DIGITAL TRANSMISSION:
SPECTRALLY EFFICIENT TECHNIQUES

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical Engineering
in the
University of Canterbury
Christchurch, New Zealand

by

B. G. Henderson B.E.(Hons)

1984

PREFACE

The use of digital signals is well established as a means of representing information. A significant advantage of this format is that once information is digitised, the means by which it can be transmitted are independent of its original form. Thus, a digital transmission system can be used to transfer, for example, voice, visual, written or computer based information. The system does not require specific knowledge of the original or final form of the information it is carrying, but it must maintain an acceptable quality in terms of the errors which can be tolerated in the received signal. This thesis is, in general, concerned with digital transmission systems and the means by which digital information can be transmitted in an efficient manner while maintaining the required quality of the received signal.

In general there will be multiple users of a transmission medium and to enable their co-existence without excessive mutual interference, the allowable spectral occupancy and/or power for each information source must be limited. The material presented in this thesis is biased towards spectrally efficient techniques of digital transmission. However, an awareness is always retained of the requirement for efficiencies in the other system variables of signalling power, complexity and cost.

The impetus for this research stemmed from an initial goal to design and construct a spectrally efficient, but relatively simple system for 9600 bits/sec duplex data-over-voice transmission on the subscriber telephone loop. Vestigial sideband (VSB) was considered as a transmission technique which met the required criteria. However, on encountering preliminary design and construction difficulties with one form of the VSB system, alternative system models were researched.

It became apparent that even though VSB is an established data transmission technique, its treatment in previous literature was sporadic and some derivatives of the fundamental VSB system had received no attention. The research goal shifted direction slightly to a study rather than a development of VSB as one alternative for spectrally efficient transmission.

The research proved to be successful in several areas:

(a) A study of VSB signalling and the alternative generation and detection techniques which could be used gives a unified reference for initial VSB system design considerations, especially if digital processing circuitry is to be employed.

(b) A method was developed for both carrier and bit synchronization of a VSB receiver which offers a simple but efficient technique of maintaining both levels of synchronization without a need for enforcing code restrictions.

(c) Even though a VSB signal does not have a constant envelope it is shown to be a possible candidate for transmission over nonlinear channels, since even after it undergoes hardlimiting it retains good spectral efficiency and suffers only a small degradation in power efficiency.

(d) VSB signalling and quadrature modulation are generally considered to be mutually exclusive methods of obtaining spectrally efficient transmission; however, a new signalling format termed quadrature VSB (QVSB) which combines the two techniques was developed and analysed. The result is a method of increasing spectral efficiency while suffering only a small degradation in power efficiency.

Papers which have resulted from the research towards this thesis are:

Henderson, B.G., and J.A. Webb, "A bandwidth-efficient data transmission system," Proc. Pacific Telecom. Conf., pp. 121-122, Jan. 1982.

Henderson, B.G., and J.A. Webb, "Synchronizing VSB data transmission," J. Electrical and Electronic Engineering, Australia. To be published.

Henderson, B.G., and J.A. Webb, "Quadrature vestigial sideband (QVSB) data transmission," submitted to IEEE Trans. Commun. Currently at the second stage of reviewing.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Mr. J.A. Webb. The ideas and time he input to my research and his unabating enthusiasm and encouragement is gratefully acknowledged. I am also grateful to Dr. P. Gough for his comments and assistance as an associate supervisor, especially during the period Mr. J.A. Webb was overseas.

I wish to thank my fellow students for their assistance, in particular Mr. A. Hamilton for his companionship and discussions on our respective research areas.

The financial support of the New Zealand Post Office is acknowledged with appreciation. Without this support, the time I dedicated to my research would not have been possible.

To my family and friends, I must express my deepest thanks for their never-ending support and encouragement.

CONTENTS

GLOSSARY

Unless otherwise stated, symbols and terms used have the following meanings:-

$\oplus$	(in equations)	exclusive-OR, modulo-2 addition
$\oplus$	(in figures)	summation
$\otimes$	(in figures)	multiplication
$*$	(on line)	convolution
$*$	(superscript)	complex conjugate
$\delta(t)$		Dirac delta function
$\delta_{k\ell} = \begin{cases} 1, & k = \ell \\ 0, & \text{elsewhere} \end{cases}$		Kronecker delta function
$\prod_{k=0}^{N} x_k = x_0 x_1 x_2 \ldots x_N$		
ρ		filter rolloff factor
σ_α^2		variance of random variable $\alpha(t)$
$\sum_{k=0}^{N} x_k = x_0 + x_1 + x_2 + \ldots + x_N$		
D		delay operator
$E[\cdot]$		Expectation value of terms inside [ ]
E_b		energy per bit
E_s		energy per symbol
$\text{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^\infty e^{-z^2} dz$		complementary error function
f		frequency variable
f_k		specific frequency

$f(\cdot)$	function of terms inside ( )
$f_i(t)$	instantaneous frequency
$F[\cdot]$	Fourier transform of terms inside [ ]
$H[\cdot]$	Hilbert transform of terms inside [ ]
$Im[\cdot]$	imaginary part of terms inside [ ]
j	$\sqrt{-1}$
$k!$	k factorial
N_o	one-sided spectral density of additive white Gaussian noise
P	average power
P_b	probability of a bit error (bit-error-rate)
P_c	probability of a symbol error
P_{de}	probability of a bit error when using differential encoding
P_e	probability of error
$p(x')$	probability density function of variable x'
$p(x'\|x)$	condition probability density function of variable x', given x has occurred
$P(A)$	probability of the occurrence of event A
$P(A\|B)$	conditional probability of occurrence event A, given event B has occurred
R	information rate in bits/sec
R_c	symbol rate in symbols/sec
$Re[\cdot]$	real part of terms inside [ ]
$R_{xx}(\tau)$	autocorrelation function of x(t)
$R_{xy}(\tau)$	cross-correlation function of x(t) and y(t)
$S_x(f)$	spectral density of X(f)
t	time variable
t_k	specific time
τ	difference between two points in time
T	bit period
T_c	symbol period

$x(t)$	function of time } Fourier transform pair
$X(f)$	function of frequency } Fourier transform pair
x_k	k^{th} sample of variable x
$\underline{x}$	matrix - referred to as a vector if a single-row matrix
$\underline{x}^T$	matrix transpose
$\underline{x}^{-1}$	matrix inverse
$\lvert\underline{x}\rvert$	determinant of a matrix
$\lvert x\rvert$	magnitude of x
$\hat{x}$	estimate of x
ω	radian frequency variable = $2\pi f$
ω_k	specific radian frequency = $2\pi f_k$
AM	amplitude modulation
APK	amplitude and phase-shift keying
ASK	amplitude-shift keying
AWGN	additive white Gaussian noise
BER	bit-error-rate
BPSK	binary phase-shift keying
BSC	binary symmetric channel
dB	decibels
d.c.	direct current
demod	demodulator
DSB	double sideband
FFT	fast Fourier transform
FM	frequency modulation
FSK	frequency-shift keying
HPA	high power amplifier
Hz	Hertz
IJF-OQPSK	intersymbol-interference and jitter-free offset quaternary phase-shift keying
ISI	intersymbol interference

mod	modulator
MSK	minimum-shift keying
OQPSK	offset quaternary phase-shift keying
PCM	pulse code modulation
PLL	phase lock loop
PM	phase modulation
PRS	partial response signalling
PSK	phase-shift keying
QORC	quadrature overlapped raised-cosine
QPSK	quaternary phase-shift keying
QVSB	quadrature vestigial sideband
SNR	signal-to-noise ratio
SQORC	staggered quadrature overlapped raised-cosine
SSB	single sideband
TWTA	travelling wave tube amplifier
VSB	vestigial sideband

ABSTRACT

Spectrally efficient methods of digital transmission are studied, with particular emphasis on vestigial sideband (VSB) signalling. A general overview of digital transmission techniques gives a basis for more detailed studies into the areas of interest. Alternative methods of generating and detecting VSB are covered, offering possible ways of simplifying system implementation. A method of synchronizing VSB is proposed and analysed, giving a simple solution to the synchronization problem. VSB, as a spectrally efficient signalling format, is shown to be suitable for transmission over nonlinear channels; its performance comparing favourably with alternative systems. A new method of attaining spectrally efficient transmission is analysed. This combines VSB and quadrature modulation to obtain a signal with double the spectral efficiency of VSB. With suitable decoding this increased spectral efficiency is obtained at the expense of only a small sacrifice in power efficiency.

CHAPTER 1

INTRODUCTION

1.1 OUTLINE OF THESIS

The main emphasis in this research towards spectrally efficient transmission techniques has been towards the vestigial sideband (VSB) signalling format. However, VSB is by no means considered in isolation, since it is just one of many techniques available for obtaining a spectrally efficient signal. Much of the analysis and results presented will have general application to the field of digital data transmission.

Prior to any analysis of VSB signalling, an overview of digital transmission techniques is given in Chapter 2. The general digital transmission model is introduced and, using this, Shannon's classical result for the capacity of a channel indicates how a compromise must be reached with the general system parameters of spectral efficiency, power efficiency, complexity and cost. Carrier-modulated systems are an extension of baseband transmission, so the stages of the general baseband model are defined as a lead-up to applying them to out-of-baseband transmission. The baseband antipodal signalling model is generally accepted as a datum for digital transmission, so this system is defined and analysed; the results of which are referenced repeatedly in the following sections and chapters.

The attractiveness of VSB as a spectrally efficient data transmission method can be somewhat negated by the difficulty in its generation and detection. The availability of low cost digital processing circuitry opens up several alternatives for the generation and detection of VSB, other than the conventional approaches, which were once considered

impractical when using analogue processing techniques. No previous studies appear to have addressed themselves to any unified study of these alternative techniques; this is the purpose of the material presented in Chapter 3.

In this treatment of digital transmission systems only synchronous transmission has been considered since asynchronous techniques cannot offer good spectral efficiency coupled with power efficiency. Bit synchronization is one level of receiver synchronization and is generally independent of the data transmission signalling format used. Another level of receiver synchronization is required in carrier-modulated systems and that is for the carrier itself. Chapter 4 is a study of VSB synchronization starting with a description of the problem followed by existing methods of synchronization and a new technique which retains power efficiency and ensures both levels of receiver synchronization are maintained. VSB synchronization poses some unique problems. If no additional synchronization information is added to the VSB signal then sufficient information may not be present in the received signal to maintain either of the two levels of synchronization (bit or carrier). This occurs for certain transmitted code sequences. Though code restriction appears to be one possible solution to maintaining synchronization, it is shown that a conflict exists between carrier and bit synchronization; any code restrictions to ensure information is always present for maintaining one level of synchronization is to the detriment of the other level of synchronization. A review of existing methods of VSB carrier synchronization where no code restriction exists shows they rely on the insertion of one or more pilot tones. The simplest of these techniques use large pilot tones, hence are power inefficient. The more power efficient of these techniques suffer from complexity and sensitivity of the recovery circuitry. A common factor with the existing carrier synchronization techniques is that none consider the problem of bit synchronization; presumably consid-

ering it as an unrelated problem. It was clear during the research on VSB that the need existed for a synchronization technique which was both simple and power efficient. Both levels of synchronization should be considered jointly, since both are important to the overall performance. Code restrictions were to be avoided, since these limit the versatility of a system. These requirements set the base for the development of a new system which involved the addition of two low level pilot tones to the VSB signal. This was achieved by simply adding a stream of synchronization pulses to the data bits prior to transmission. Rather than rely solely on the pilots for synchronization, as is the general situation with existing systems, the pilots in the new system are used as a secondary source of information, whereby their presence become of importance only when the data signal it - self does not contain synchronization information. The performance of this new system is derived and, using computer simulation, estimates of carrier phase and bit timing variations are obtained.

The performance of various signalling formats when operating over a nonlinear channel has received considerable attention over the past decade, with particular emphasis on their application to satellite communication. This is the subject of Chapter 5. If the power source at a repeater station such as a satellite is limited, then its efficient operation, in terms of keeping its output signal power above extraneous noise levels, may require it to operate in a nonlinear mode. The presence of a nonlinear repeater, typically a travelling wave tube amplifier (TWTA) or a high power amplifier (HPA), introduced AM-AM and AM-PM distortion if the repeater is operated in a power efficient mode that is near or at saturation level. The conventional approach at arriving at a modulation format suitable for transmission over such a nonlinear channel has therefore been to use a constant envelope signal which does not suffer the AM-AM or AM-PM distortions. However, as shown in Chapter 5, such signals are generally

inefficient in terms of spectral occupancy and additional filtering is required to increase this efficiency. A review of existing modulation formats for nonlinear transmission is presented and it is shown how a signal's instantaneous phase and envelope variations can be used to predict and analyse the effect of filtering and nonlinear amplification. Since filtering is seen to introduce envelope variations, the question was posed; if filtering a constant envelope signal to increase its spectral efficiency introduces envelope variations, why not use a signalling format which is initially spectrally efficient and contains envelope variations in its original form? This led to a study of VSB as a possible candidate for transmission over a nonlinear channel. An equivalence between VSB and filtered offset quaternary phase shift keying (OQPSK), derived in Chapter 3, enables the results for VSB to be applied also to OQPSK. Methods of obtaining spectral density, envelope, phase and eye-pattern plots using baseband models are derived. This simplifies the computer simulation of the nonlinear VSB system. The plots are then used to obtain performance figures for VSB in terms of spectral efficiency and degradation, which in turn are used in a comparison of VSB with existing systems used over nonlinear channels. VSB is seen to be a relatively robust and spectrally efficient format for nonlinear transmission.

Since a VSB signal consists of quadrature related components it has generally not been considered as a suitable signalling format for combining with quadrature modulation to further increase its spectral efficiency. However, in Chapter 6 such a combination is proposed and analysed. The new signalling technique is given the name quadrature VSB (QVSB). The similarity between QVSB and partial response signalling (PRS) is demonstrated. The distinguishing feature of QVSB is that the controlled intersymbol interference (ISI) occurs between the quadrature related signal channels as opposed to existing PRS systems where the ISI is kept within the one phase plane. The similarity between QVSB and PRS systems is used

in deriving suitable encoding/decoding techniques to "unwind" the controlled ISI in the QVSB signal. As with PRS systems, the precoding technique offers the simplest method, but also the most inefficient in terms of bit-error-rate performance. A significant improvement in performance is seen to be gained by employing probabilistic decoding techniques. The Viterbi algorithm, which has been applied to convolutional coded and PRS systems, is applied to QVSB. The Viterbi algorithm is generally applied on the assumption that noise samples input to the decoder are statistically independent and this is used as a first step in deriving a probabilistic decoder. The optimum decoder is then derived where the Viterbi algorithm is again used but this time the noise statistics are taken into account. QVSB is shown to offer a 2 to 3 dB signal-to-noise power advantage over conventional systems with comparable spectral efficiency. An additional and very interesting observation made during the study of QVSB was the use of the absolute value operation to replace squaring in probabilistic decoding algorithms. This approximation can be used to significantly simplify the computational effort required while apparently introducing only a relatively small performance degradation.

1.2 USE OF "IDEAL" MODELS AND COMPUTER SIMULATION

For the majority of the analysis presented in the following chapters, the digital transmission system models are "ideal", in that they have optimum signal passband shaping, perfect receiver synchronization and specifically defined channel impairments. This is by no means the situation in an actual system, but the restriction of model parameters to the "ideal" not only simplifies analysis, it enables a comparison with other systems treated in a similar manner which would be difficult if non-standard signal impairments were employed. Though the performance of actual systems will inevitably be worse than their "ideal" counterparts, an indication of the

absolute and relative performances of the various systems can be obtained from the "ideal" models.

Computer simulation has been used in Chapters 4, 5 and 6 to study the performance of various systems. Its use enabled system variables to be readily altered and produced the required results without excessively complicated mathematical analysis or time consuming hardware construction. With all of the computer simulation models, the signal passband shaping filters used had gradual, hence realisable, spectral rolloff characteristics. The models and results can therefore be related to practicable systems.

1.3 COMMONLY USED CONVENTIONS AND TERMINOLOGY

The information bits and coded symbols at the transmitter are always assumed to be in the form of impulses. This is for mathematical convenience, since the impulses will simply act as time shifting and amplitude scaling functions and symbol shaping can be completely defined using the transmitter and receiver filters. This is similar to the practicable set-up where a transmitter filter's time response is stored in a transversal filter or a look-up table in a microprocessor and is clocked out every symbol period with the required amplitude and polarity.

Changing between the frequency and time domains is freely used in analysing signal characteristics. Specific mention of this is not always given, but the following convention is always adhered to. Any function defined in the time domain using lower-case notation x(t), say, will have its Fourier transform defined in the frequency domain using upper-case notation X(f), where

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft} df. \qquad (1.1)$$

Also, a change is freely made between the variables ω in radians/sec, and f in Hz, where

$$\omega = 2\pi f. \qquad (1.2)$$

CHAPTER 2

OVERVIEW OF DIGITAL TRANSMISSION TECHNIQUES

2.1 INTRODUCTION

Using a general model for a digital system, the parameters commonly used in comparing different transmission methods will be defined. A theoretical bound for these parameters is given by Shannon's classical result for the capacity of a channel. While this bound cannot, as yet, be attained in practice, Shannon does indicate the means by which it can be approached.

Most of the discussion in the following chapters is concerned with systems where modulation or frequency translation are used to obtain out-of-baseband data signals in the channel. Such systems are an extension of the general baseband model, so the various stages and some specific forms of this model are studied before extending it to include the out-of-baseband systems.

2.2 GENERAL MODEL FOR A DIGITAL TRANSMISSION SYSTEM

The fundamental stages of a digital transmission system are shown in Fig. 2.1. The distinguishing feature of this system is that the input and output information is digital, usually in the form of binary digits (or bits). The actual signal levels used for the information bit stream $d(t)$ will be assumed to be +1 and -1 corresponding, respectively, to binary levels "1" and "0". Also, the mathematical treatment of the system is simplified if the information bits are considered in impulse form:

$$d(t) = \sum_{k=0}^{N} d_k \, \delta(t - kT), \tag{2.1}$$

where $d_k = \pm 1$. T is the bit period, hence the information rate is

$$R = \frac{1}{T} \text{ bits/sec.} \tag{2.2}$$

It is often convenient to consider the bit stream in vector form:

$$\underline{d} = (d_0, d_1, d_2, \ldots, d_k, \ldots, d_N). \tag{2.3}$$

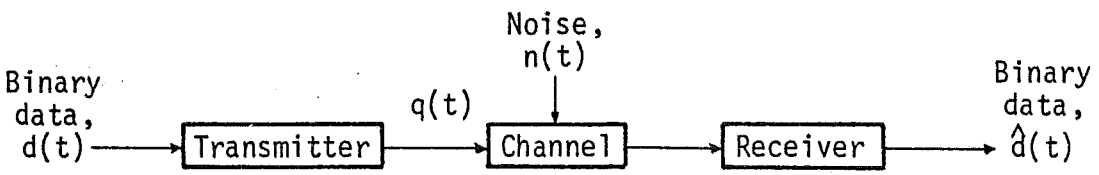


Fig. 2.1 General model of a digital transmission system.

The transmitter's purpose is to change, if necessary, the information bits into a form suitable for transmission through the channel. En route to the receiver the transmitted information q(t) is inevitably perturbed by channel impairments. The receiver must then make an acceptable estimate $\hat{d}(t)$ (or in vector form $\underline{\hat{d}}$) of the information bits input to the transmitter. The number of errors in $\hat{d}(t)$ is the main criterion in assessing a digital system's quality of performance. It is measured in terms of P_b, the probability of a bit error, also referred to as the bit-error-rate (BER), where

$$P_b = \frac{\text{Number of Bit Errors in } \hat{d}(t)}{\text{Total Number of Information Bits}} \tag{2.4}$$

As examples of P_b, voice traffic when transmitted in digital form can tolerate a BER as high as 10^{-3} [Newcombe and Pasupathy, 1982]. For digital radio the upper limit on P_b is about 10^{-4} [Anderson and Barber, 1979; Oetting, 1979]. Van Trees et al. [1977] indicate a range 10^{-8} - 10^{-12} is required for computer communications or electronic mail.

Numerous transmitter and receiver configurations are available, so from an engineering point of view it is important to have parameters available for comparing the alternatives and making a choice on a system most suited to a particular application.

2.2.1 Basis for Comparing Different Systems

In trying to achieve the desired BER the main factors requiring attention are:

(a) information rate
(b) channel impairments
(c) signalling level (power)
(d) signalling bandwidth
(e) complexity
(f) cost

The usual situation is with (a) and (b) fixed, and (c) through (f) requiring minimisation. As will be shown a compromise must be reached with these latter factors.

In defining the parameters for comparing various systems it will be assumed the information rate is set by the information source independently of the transmitter and receiver configurations. Also, some standard form of channel impairment must be adopted. The most common form is the channel having uniform amplitude and phase characteristics over the signalling passband of q(t), with the only impairment as additive

white Gaussian noise (AWGN) which has a two-sided spectral density of $N_o/2$ watts/Hz.

Often a channel will have non-uniform amplitude and phase characteristics which may distort the transmitted signal. These impairments usually cause distortion in a predictable manner and can, in theory, be removed by suitable signal conditioning at the transmitter and/or receiver. If possible, the transmitted signal would be positioned within the channel passband where such distortion is minimal. Bennett and Davey [1965] have considered these impairments in comparing various systems.

Clark [1976] justifies the AWGN channel model in stating that the relative tolerance of systems to AWGN is a good indication of their relative tolerance to additive noise in general. Other important forms of additive noise are impulse noise [Bennett and Davey, 1965; Parson and Sheikh, 1979] and crosstalk from other transmission systems [Bellamy, 1982].

The fading channel is another form of channel impairment which has received considerable attention [Feher, 1981 and references therein; Wozencraft and Jacobs, 1965]. With this, the interference is assumed to comprise of both multiplicative and additive terms.

Since the AWGN model is relatively easy to handle mathematically, and so has received widespread use, it will be used as the channel impairment when comparing different systems. An exception to this is in Chapter 5 when the nonlinear channel is considered. If a repeater is included in a channel link, and its power source is limited, it may have to operate its amplification stage in a nonlinear mode to make efficient use of the available power.

A standard measure of the signalling level is the average energy per bit (E_b) at the input to the receiver. If P is the average signal

power at this point then

$$E_b = P/R \text{ Joules/bit} \qquad (2.5)$$

Clearly, the larger the ratio E_b/N_o the smaller P_b will be for a given system. Since E_b is limited by equipment capabilities and allowable crosstalk into other systems it is desirable to minimise this value while still maintaining the desired P_b.

Minimising the signalling bandwidth allows efficient use to be made of the available transmission medium. If a channel bandwidth of W Hz is required for the transmission of R bits/sec, the ratio R/W bits/sec/Hz is used to define a system's spectral efficiency.

Unfortunately, as shown in section 2.2.2, a compromise is required between the spectral efficiency and power efficiency of a system. Both of these resources can, however, be traded for a third, namely, complexity. It is really only since the mass production of large scale integrated circuitry this resource can be utilised without excessive cost. Increasing the "intelligence" of the transmitter and/or receiver allows some freedom in the choice of signalling power and bandwidth. Systems efficient in both of these latter two areas, which were once considered impractical due to high costs, may now be feasible to implement.

2.2.2 Upper Bound in Performance

The classical result for the capacity of a channel in the presence of AWGN was first proved and stated by Shannon in 1948. Quoting a theorem from one of his later papers, Shannon [1949] states:

> "Let P be the average transmitter power and suppose the noise is white thermal noise of power N in the band W. By sufficiently complicated encoding systems it is possible to transmit binary digits at a rate
>
> $$C = W \log_2\left(\frac{P+N}{N}\right) \tag{2.6}$$
>
> with as small a frequency of errors as desired. It is not possible by any encoding method to send at a higher rate with an arbitrarily low frequency of errors."

The units of C are bits/sec. Using the substitutions $P = CE_b$ and $N = N_oW$, (2.6) can be rewritten in the form

$$C/W = \log_2\left(1 + \frac{CE_b}{WN_o}\right) \text{ bits/sec/Hz}, \tag{2.7}$$

and from this,

$$E_b/N_o = \frac{2^{C/W} - 1}{C/W} . \tag{2.8}$$

Equation (2.8) gives the required E_b/N_o for obtaining the theoretical maximum spectral efficiency of C/W. For any practicable system R/W will be less than C/W.

Using (2.8), the relationship between E_b/N_o and C/W has been plotted in Fig. 2.2. This illustrates the significance of Shannon's theorem. Any practicable system, where R rather than C is of interest, will have its operating point above and to the left of the curve in Fig. 2.2. The complexity of a particular system will increase as its operating point moves closer to the curve. For fixed information rate and complexity the trade off between power and bandwidth is evident. If a small E_b/N_o is desired, then W must be increased, so reducing the

spectral efficiency. The limiting value of E_b/N_o is found when infinite bandwidth is available, i.e., C/W = 0. Substituting C/W = 0 into (2.8) gives E_b/N_o = -1.6 dB (Appendix 1). If E_b/N_o is less than this value successful transmission is impossible. In the alternative direction, for a fixed complexity, increasing a system's spectral efficiency requires increasing E_b/N_o.

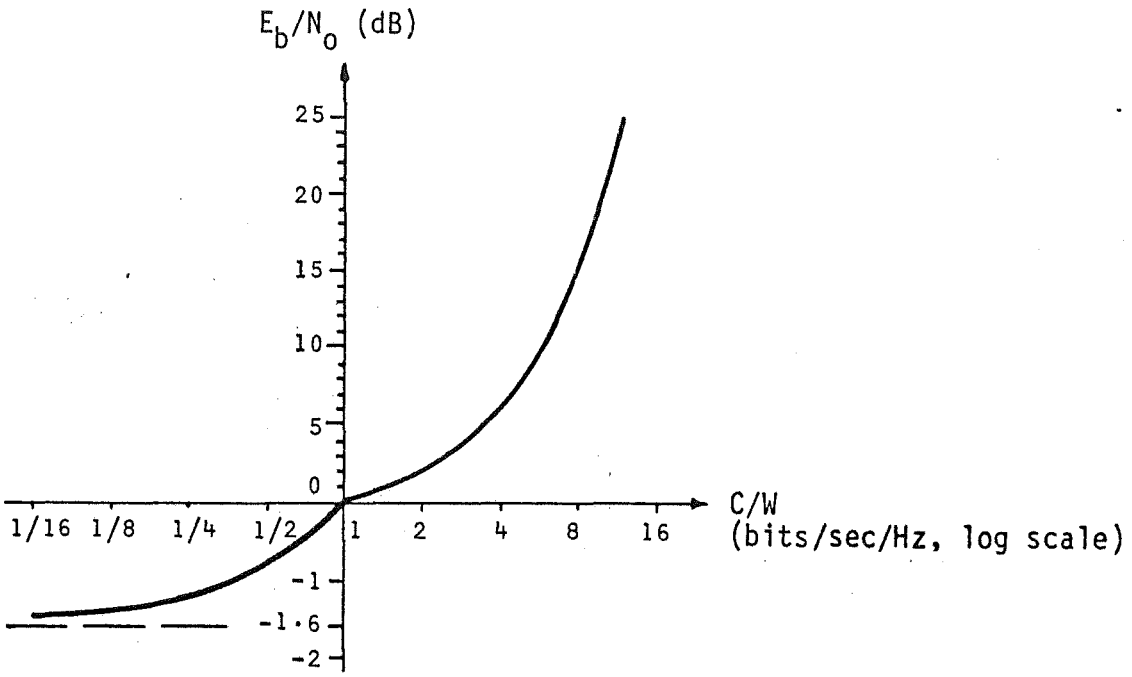


Fig. 2.2 E_b/N_o versus C/W. Note change in E_b/N_o scale at origin.

2.3 PRINCIPLES OF BASEBAND TRANSMISSION

A general model for a baseband data system is shown in Fig. 2.3. The encoder and decoder are not necessarily linear, but it will be assumed the transmitter and receiver filters, respectively, $g_1(t)$ and $g_2(t)$ are linear with

$$g(t) = g_1(t) * g_2(t). \tag{2.9}$$

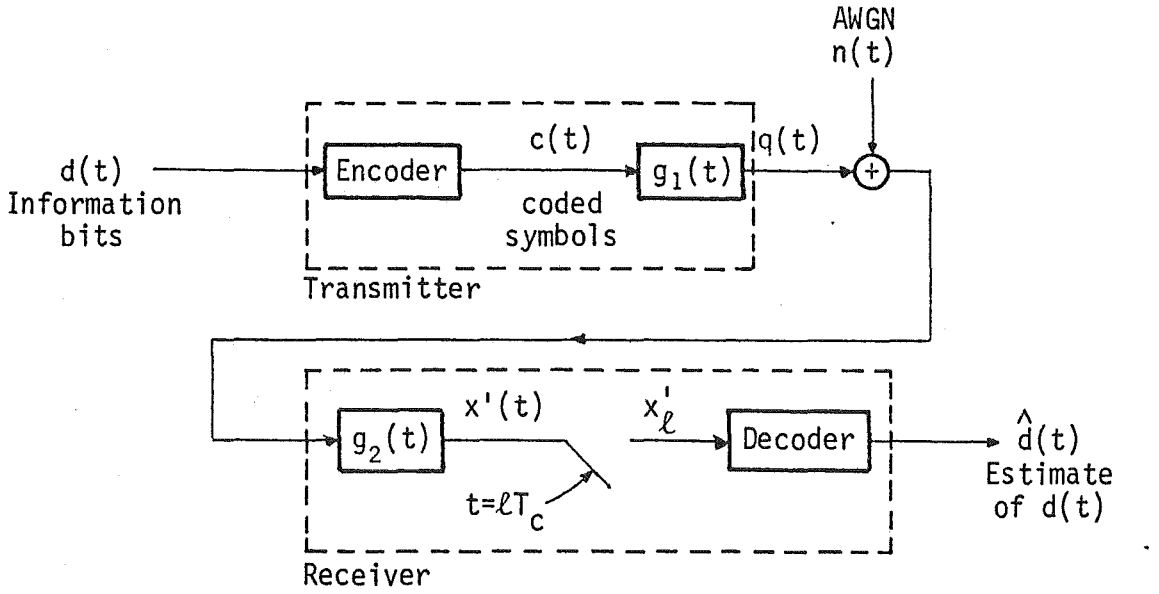


Fig. 2.3 Stages of a baseband system.

2.3.1 Encoding

The encoder transforms the information bits into discrete symbols to achieve increased transmission efficiency. The coded symbols can be defined as

$$c(t) = \sum_{\ell=0}^{M} c_\ell \, \delta(t - \ell T_c), \tag{2.10}$$

or in vector form

$$\underline{c} = (c_0, c_1, c_2, \ldots, c_\ell, \ldots, c_M). \tag{2.11}$$

With the symbol period as T_c, the symbol rate will be $R_c = 1/T_c$. The ratio R/R_c gives the number of information bits per transmitted symbol. In general, $R/R_c > 1$ for spectrally efficient systems and $R/R_c < 1$ for power efficient systems.

With systems such as partial response signalling (PRS), the

encoding operation may be performed by $g_1(t)$ and possibly $g_2(t)$. In this case R_c is defined at the output of $g_1(t)$.

2.3.2 Spectral Efficiency and E_b

In Fig. 2.3, $g_1(t)$ determines the shaping of the transmitted symbols. This filter governs the spectral characteristics, hence spectral efficiency of the channel signal $q(t)$ and, in conjunction with symbol sequence $\underline{c}$, controls the signalling level in the channel.

Assuming $c(t)$ consists of random data, its autocorrelation function is (see Appendix 2)

$$R_{cc}(\tau) = \frac{E[c_\ell^2]\ \delta(\tau)}{T_c} . \tag{2.12}$$

Using (2.12), the spectral density function for $c(t)$ is

$$\begin{aligned} S_c(f) &= F[R_{cc}(\tau)] \\ &= \frac{E[c_\ell^2]}{T_c} , \end{aligned} \tag{2.13}$$

which leads to the spectral density of $q(t)$;

$$\begin{aligned} S_q(f) &= S_c(f)|G_1(f)|^2 \\ &= \frac{E[c_\ell^2]}{T_c} |G_1(f)|^2 . \end{aligned} \tag{2.14}$$

If $S_q(f)$ has significant magnitude up to some frequency W_q Hz, the spectral efficiency of the system is R/W_q bits/sec/Hz.

The area under $S_q(f)$ gives the average power of q(t). The average energy per symbol is therefore given by

$$E_s = T_c \int_{-\infty}^{\infty} S_q(f)\, df \quad \text{Joules/symbol.} \tag{2.15}$$

Substituting (2.14) into (2.15) gives

$$E_s = E[c_\ell^2] \int_{-\infty}^{\infty} |G_1(f)|^2 df. \tag{2.16}$$

E_b can now be found using (2.16) and the relationship

$$E_b = E_s \frac{R_c}{R}. \tag{2.17}$$

2.3.3 Optimum Receiver Filtering

The role of $g_2(t)$ in the system model is to maximise the SNR at its output. In addition to $g_1(t)$, $g_2(t)$ usually modifies the shape of the symbols prior to sampling and decoding. This latter point is considered in section 2.3.5.

For a received signal of the form

$$r(t) = \sum_{\ell=0}^{M} c_\ell g_1(t - \ell T_c) + n(t), \tag{2.18}$$

where $g_1(t)$ is square-integrable and n(t) is AWGN, the optimum form of $g_2(t)$ is [Forney, 1972; Wozencraft and Jacobs, 1965]

$$g_2(t) = g_1(-t), \tag{2.19a}$$

or equivalently in the frequency domain,

$$G_2(f) = G_1^*(f). \tag{2.19b}$$

The output of $g_2(t)$ sampled at a rate R_c forms a set of sufficient statistics for estimating $\underline{c}$ and hence $\underline{d}$. This is known as matched filtering.

An equivalent operation to the matched filter is the correlation receiver [Haykin, 1978]. With this latter method each symbol in the received signal $r(t)$ is multiplied by a stored replica of the symbol shape. The product is integrated over the duration of the received symbol, after which time the integrator output is sampled and processed by the decoder as for matched filtering.

2.3.4 Optimum Decoding

The desired signal at the output of $g_2(t)$ is

$$x(t) = c(t) * g(t). \tag{2.20}$$

This is, however, perturbed by the filtered AWGN

$$\alpha(t) = n(t) * g_2(t), \tag{2.21}$$

giving the composite signal

$$x'(t) = x(t) + \alpha(t). \tag{2.22}$$

The noise $\alpha(t)$ will be Gaussian with autocorrelation function

$$R_{\alpha\alpha}(\tau) = \frac{N_o}{2} F[|G_2(f)|^2]. \tag{2.23}$$

With matched filtering,

$$|G_2(f)|^2 = G_1(f)G_2(f)$$

$$= G(f), \tag{2.24}$$

therefore

$$R_{\alpha\alpha}(\tau) = \frac{N_o}{2} g(\tau), \tag{2.25}$$

or in normalised form

$$R'_{\alpha\alpha}(\tau) = R_{\alpha\alpha}(\tau)/R_{\alpha\alpha}(0)$$

$$= g(\tau). \tag{2.26}$$

An optimum decoder will make use of the noise statistics given by (2.26) in arriving at its estimate.

The signal input to the decoder is in discrete samples

$$x'_\ell = x_\ell + \alpha_\ell, \tag{2.27}$$

where the sequence of samples can be considered in vector form

$$\underline{x}' = \underline{x} + \underline{\alpha}. \tag{2.28}$$

An optimum decoder will have *a priori* knowledge of the noise statistics for $\underline{\alpha}$ (given by (2.26)) and of the permissible alternatives $\underline{x}$ can have. Using this and its input $\underline{x}'$, it must decide on which of the alternative possibilities of $\underline{x}$ is the most likely to have been received. The decoder then sets its estimate $\hat{\underline{x}}$ to equal this most likely alternative. It will also need to know the relationships between information bits $\underline{d}$,

coded symbols $\underline{c}$ and samples $\underline{x}$ to relate $\hat{\underline{x}}$ to an estimate $\hat{\underline{d}}$.

Wozencraft and Jacobs [1965] show that for a received vector $\underline{x}'$, the optimum decoder is one which sets estimate $\hat{\underline{x}}$ equal to the hypothesis $\underline{x}$ which maximises the conditional probability function

$$P(\underline{x}|\underline{x}') = \frac{P(\underline{x})p(\underline{x}'|x)}{p(\underline{x}')} \tag{2.29}$$

This is termed maximum *a posteriori* probability decoding. Since $\underline{x}'$ is fixed and independent of the $\underline{x}$ alternative chosen, maximising (2.29) is equivalent to maximising

$$P(\underline{x})p(\underline{x}'|\underline{x}). \tag{2.30}$$

If $\hat{\underline{x}}$ is determined using the term $p(\underline{x}'|\underline{x})$, without regard for factor $P(\underline{x})$ in (2.30), the operation is known as maximum-likelihood (ML) decoding. If each possible alternative of $\underline{x}$ is equally likely, then ML decoding will yield the minimum probability of error in estimate $\hat{\underline{x}}$.

Using variable transformation, maximising $p(\underline{x}'|\underline{x})$ is equivalent to maximising $p(\underline{x}' - \underline{x}|\underline{x})$. From (2.28), the term $(\underline{x}' - \underline{x})$ is simply the Gaussian noise which is independent of $\underline{x}$, so

$$\begin{aligned} p(\underline{x}' - \underline{x}|\underline{x}) &= p(\underline{\alpha}|\underline{x}) \\ &= p(\underline{\alpha}). \end{aligned} \tag{2.31}$$

For the special case where the components of $\underline{\alpha}$ are statistically independent, i.e.,

$$R'_{\alpha\alpha}(\ell T_c) = \begin{cases} 1, & \ell = 0 \\ 0, & \ell \neq 0, \end{cases} \tag{2.32}$$

then

$$p(\underline{\alpha}) = \frac{1}{(2\pi\ \sigma_\alpha^2)^{(M+1)/2}} \exp\left[\sum_{\ell=0}^{M} \left(\frac{-\alpha_\ell^2}{2\sigma_\alpha^2}\right)\right], \qquad (2.33)$$

where σ_α^2 is the variance or power of the noise $\alpha(t)$. Equ. (2.33) can be simplified by using the log-likelihood representation of the ML decoding operation. Since the logarithm is a monotonic function of its (always positive) argument, maximising (2.33) is equivalent to minimising

$$-\log_e[p(\underline{\alpha})] = -K_1 + K_2 \sum_{\ell=0}^{M} \alpha_\ell^2 \ . \qquad (2.34)$$

K_1 and K_2 are positive constants representing the terms independent of $\underline{\alpha}$, therefore can be ignored. Thus, the ML decoder is required to minimise, with respect to hypothesis $\underline{x}$, the function

$$J(\underline{x}) = \sum_{\ell=0}^{M} \alpha_\ell^2$$

$$= \sum_{\ell=0}^{M} (x'_\ell - x_\ell)^2 \qquad (2.35)$$

This operation minimises the Euclidean distance between $\underline{x}'$ and $\underline{x}$.

2.3.5 Symbol Shaping, Zero ISI and Eye-Patterns

The role of the transmitter filter $g_1(t)$ in determining the spectral efficiency of the signal in the channel was considered in section 2.3.2. It has an additional role, in conjunction with $g_2(t)$, in governing the shaping of each symbol prior to sampling and decoding.

A filtering condition generally desired is where each sampled signal level x_ℓ at the decoder input is directly proportioned to one, and only one, transmitted symbol. This condition is termed zero inter-

symbol interference (ISI). Original work in this area has been credited to Nyquist [1928]. Of interest here is Nyquist's first criterion for zero ISI which he derived for the case where the overall symbol shaping function G(f) has a restricted passband. A more general criterion for zero ISI was given by Gibby and Smith [1965] and requires G(f) to obey the relationships

$$\sum_{n=-\infty}^{\infty} \mathrm{Re}[G(\beta + nR_c)] = K/R_c, \quad K = \text{real constant} \tag{2.36a}$$

and

$$\sum_{n=-\infty}^{\infty} \mathrm{Im}[G(\beta + nR_c)] = 0, \tag{2.36b}$$

where R_c equals the symbol, hence sampling rate, and β with units of Hz has a range

$$-R_c/2 \leq \beta \leq R_c/2.$$

Equ. (2.36) is equivalent to requiring

$$g(\ell T_c) = K\delta_{\ell 0}. \tag{2.37}$$

With matched filtering, G(f) will have no imaginary part. Also, if both $G_1(f)$ and $G_2(f)$ are real, then

$$G_1(f) = G_2(f), \tag{2.38}$$

hence

$$g_1(t) = g_2(t). \tag{2.39}$$

With this situation, two extremes in $g_1(t)$ and $g_2(t)$ are possible. These are illustrated in Fig. 2.4. The rectangular pulse in Fig. 2.4(a) is usually set to extend for the duration of the symbol period, with $\tau = T_c$. It has an advantage of being easily generated, however, as shown by its Fourier transform, it has poor spectral efficiency with the slowly decaying rolloff. The rectangular passband in Fig. 2.4(b) offers maximum spectral efficiency, but has a disadvantage that its realisation would require an infinite time delay [Sunde, May 1954].

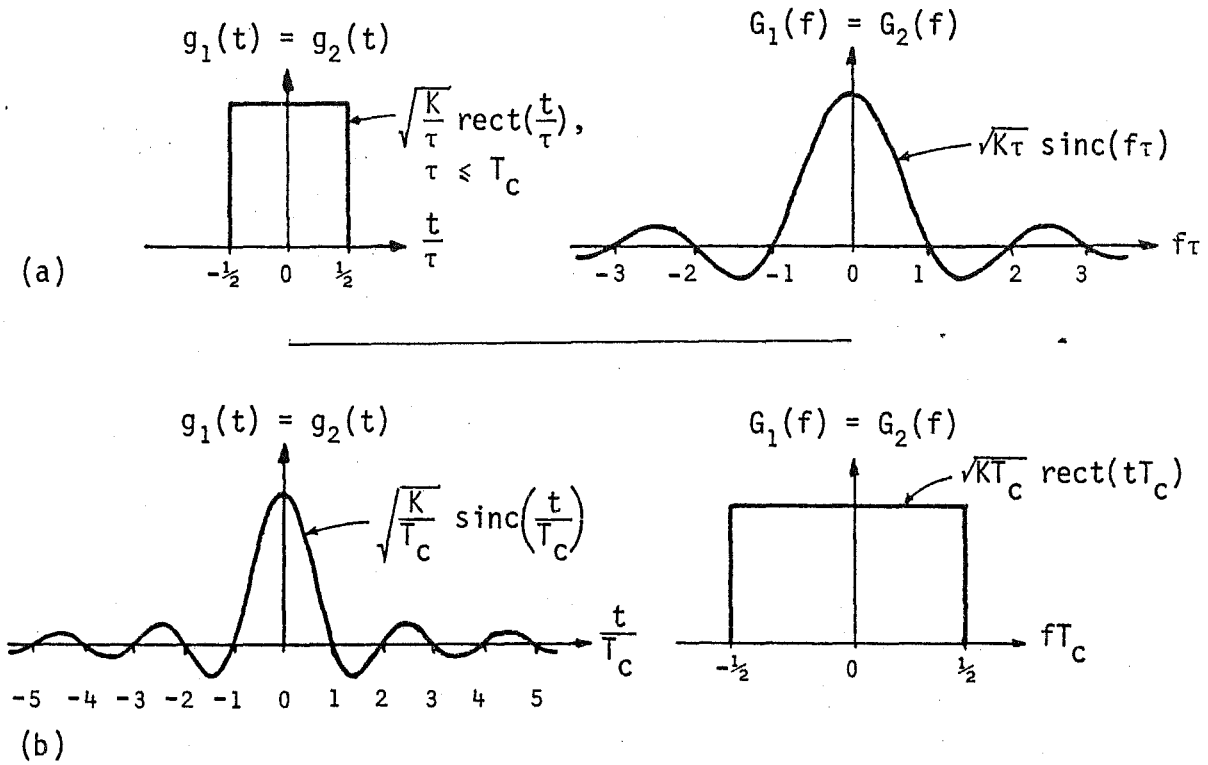


Fig. 2.4 Two extremes in signal shaping which confirm with zero ISI criterion. Fourier transform pair shown for each. (a) Rectangular pulse. (b) Rectangular passband.

Symbol shaping for practicable systems will be somewhere between the two extremes in Fig. 2.4. One such configuration is identical cosine rolloff filters used for $G_1(f)$ and $G_2(f)$. Their cascade results in G(f) having a raised-cosine spectrum. The raised-cosine filter can be defined

as follows [Haykin, 1978]:

$$U(f) = \begin{cases} \frac{1}{2B} & , \quad |f| < f_u \\ \frac{1}{4B}\left[1 + \cos\left\{\frac{\pi(|f| - f_u)}{2B - 2f_u}\right\}\right], & f_u < |f| < 2B-f_u \\ 0 & , \quad |f| > 2B-f_u. \end{cases} \tag{2.40}$$

The value B Hz in (2.40) gives the -6 dB points in the filter's rolloff. U(f) can have different rates of rolloff defined by the rolloff factor

$$\rho = 1 - \frac{f_u}{B}. \tag{2.41}$$

The time response of the raised-cosine filter is

$$u(t) = F[U(f)]$$

$$= \frac{\text{sinc}(2Bt)\cos(2\pi\rho Bt)}{1 - 16\rho^2B^2t^2}. \tag{2.42}$$

Fig. 2.5 shows examples of U(f) and u(t) for different ρ values. The unrealisable rectangular passband results when $\rho = 0$. For zero ISI transmission using raised-cosine filtering, then

$$G(f) = K\ U(f), \tag{2.43}$$

with

$$B = \frac{1}{2T_c}. \tag{2.44}$$

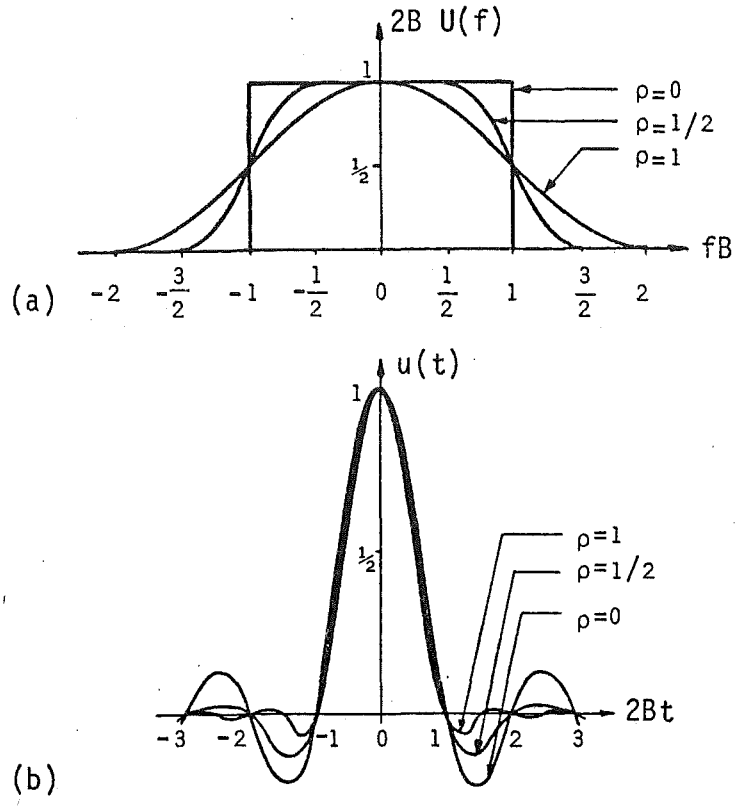


Fig. 2.5 Raised-cosine rolloff filters. (a) Frequency response. (b) Time response.

A visual indication of the extent of ISI is the eye-pattern. This is obtained by displaying, on say an oscilloscope, each symbol period of x(t) overlapped on the same time base. An example of this is shown in Fig. 2.6 for the case of random data, two-level symbols and with G(f) a raised-cosine filter with $\rho = 0.5$. Normally a display of the received signal without noise is desired so the signal distortion due to filtering can be observed. Fig. 2.6 illustrates that there is only one point in time during each symbol period where x(t) will be at one of two specific levels. This is the point of zero ISI at which sampling should occur if maximum noise immunity is to be achieved.

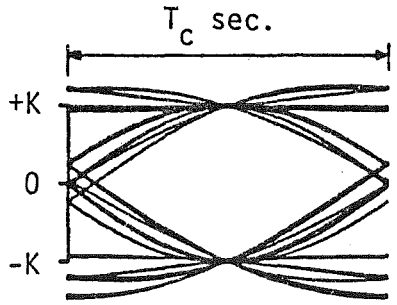


Fig. 2.6 Two level eye-pattern. Computer simulated using a raised-cosine passband with $\rho = 0.5$

The amplitude of the eye opening is an indication of a system's margin over noise and its width indicates the allowable timing error in sampling the signal. Van den Elzen [1975] gives a detailed analysis of the worst-case eye openings for several data transmission systems. Smith [1982] studies the distortion of eye openings for the particular case of a raised-cosine shaping filter when timing and carrier phase errors in carrier-modulated systems are present.

2.4 BASEBAND ANTIPODAL SIGNALLING OVER A BINARY SYMMETRIC CHANNEL

The performance of the baseband system which uses antipodal signalling over a binary symmetrical channel (BSC) is generally used as a datum against which other transmission systems are compared. A BSC can be described by the transition diagram shown in Fig. 2.7 [Wozencraft and Jacobs, 1965]. It has no memory and transmits and receives two symbols. Correct reception occurs if b_1 is received when a_1 is transmitted, or if b_2 is received when a_2 is transmitted. This has a probability $(1 - P_c)$, where P_c is the probability of an error. An error occurs if b_1 is received when a_2 is transmitted, or b_2 is received when a_1 is transmitted. In the case of antipodal signalling $a_1 = -a_2$ and $b_1 = -b_2$.

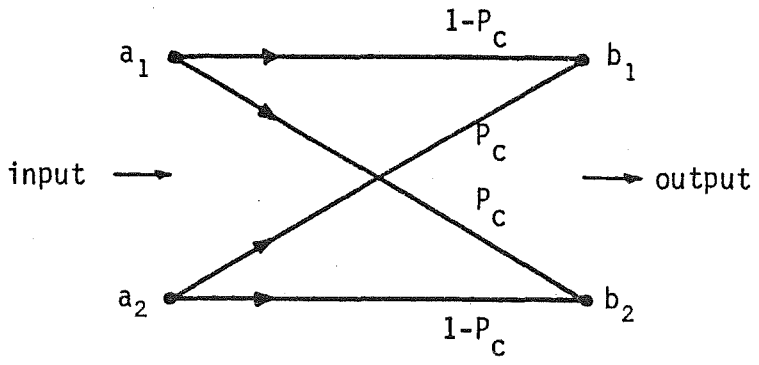


Fig. 2.7 Transition diagram for a BSC.

Of interest here is the performance of the baseband system where no encoding is used; i.e.,

$$c_k = d_k$$

$$= \pm 1, \tag{2.45}$$

and

$$T_c = T. \tag{2.46}$$

Zero ISI requires

$$x_k = Kd_k, \tag{2.47}$$

where the time delay between symbol transmission (an information bit in this case) and reception has been ignored since this time is fixed.

The BER performance of the system will be derived assuming AWGN in the channel and matched filtering, with zero ISI. The correlation among noise samples $\underline{\alpha}$ is therefore determined by g(t) (see (2.26)), and since zero ISI exists the noise samples must be statistically independent. The decoding requirements for this particular case were considered at the

end of section 2.3.4. With zero ISI the signal samples $\underline{x}$ are also statistically independent, so the optimum decoder can make its estimates on a bit-by-bit basis. Thus, for each sample the decoder makes the decision

$$\hat{d}_k = \begin{cases} +1, & x_k' > 0 \\ -1, & x_k' < 0. \end{cases} \tag{2.48}$$

With $\underline{d}$ consisting of a random sequence, the probability of a bit error is

$$P_b = \tfrac{1}{2}p(x_k' < 0 | d_k = 1) + \tfrac{1}{2}p(x_k' > 0 | d_k = -1)$$

$$= p(x_k' < 0 | d_k = 1). \tag{2.49}$$

x_k' consists of the Gaussian noise term α_k, with zero mean and variance σ_α^2, which is offset by the desired signal x_k. The probability density function of x_k' is therefore

$$p(x_k') = \frac{1}{\sigma_\alpha \sqrt{2\pi}} \exp\left[\frac{-(x_k' - x_k)^2}{2\sigma_\alpha^2}\right]. \tag{2.50}$$

Using (2.50), (2.49) becomes

$$P_b = \frac{1}{\sigma_\alpha \sqrt{2\pi}} \int_{-\infty}^{0} \exp\left[\frac{-(x_k' - K)^2}{2\sigma_\alpha^2}\right] dx_k', \tag{2.51}$$

where K is given in (2.47). Making the substitution

$$\gamma = \frac{(x_k' - K)}{\sqrt{2}\,\sigma_\alpha}, \tag{2.52}$$

(2.51) reduces to

$$P_b = \frac{1}{\sqrt{\pi}} \int_{\frac{K}{2\sqrt{\sigma_\alpha}}}^{\infty} \exp(-\gamma^2)d\gamma$$

$$= \frac{1}{2}\,\mathrm{erfc}\left(\frac{K}{\sqrt{2}\,\sigma_\alpha}\right), \tag{2.53}$$

where erfc(·) is known as the complementary error function [Haykin, 1978].

From (2.37),

$$K = g(0)$$

$$= \int_{-\infty}^{\infty} \delta(t)[g_1(t) * g_2(t)]dt$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t)g_1(\tau)g_2(t-\tau)d\tau dt, \tag{2.54}$$

and since matched filtering is assumed, (2.54) reduces to

$$K = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t)g_1(\tau)g_1(\tau - t)dtd\tau$$

$$= \int_{-\infty}^{\infty} g_1^2(\tau)d\tau. \tag{2.55}$$

Since $g_1(t)$ is real, Rayleigh's energy theorem [Haykin, 1978] enables (2.55) to be rewritten as

$$K = \int_{-\infty}^{\infty} |G_1(f)|^2 df. \tag{2.56}$$

The variance or power of the noise samples is given by

$$\sigma_\alpha^2 = \frac{N_o}{2}\int_{-\infty}^{\infty} |G_2(f)|^2 df$$

$$= \frac{N_o}{2}\int_{-\infty}^{\infty} |G_1(f)|^2 df$$

$$= \frac{N_o}{2} K. \qquad (2.57)$$

Calculation of E_b was considered in section 2.3.2. By using (2.56) the relationship between E_b and K becomes simply

$$E_b = K. \qquad (2.58)$$

Substituting (2.57) and (2.58) into (2.53) gives

$$P_b = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_o}}\right). \qquad (2.59)$$

The BER curve given by (2.59) is shown in Fig. 2.8.

In arriving at (2.59) the only restrictions placed on $g_1(t)$ and $g_2(t)$ were that they are matched and conform with the zero ISI criterion. Any of the filter shapes discussed in section 2.3.5 will, under ideal conditions, give a BER performance as shown in Fig. 2.8. Maximum spectral efficiency of 2 bits/sec/Hz would be achieved with the rectangular passband shown in Fig. 2.4(b). However, in practice, when realisable filters must be used, the spectral efficiency will be less than 2 bits/sec/Hz.

2.5 CODING FOR INCREASED EFFICIENCY

Coding of the information bits using the encoding stage and in some cases the signal passband g(t) (see Fig. 2.3) can be used to increase power efficiency or spectral efficiency. Using the simple anti-

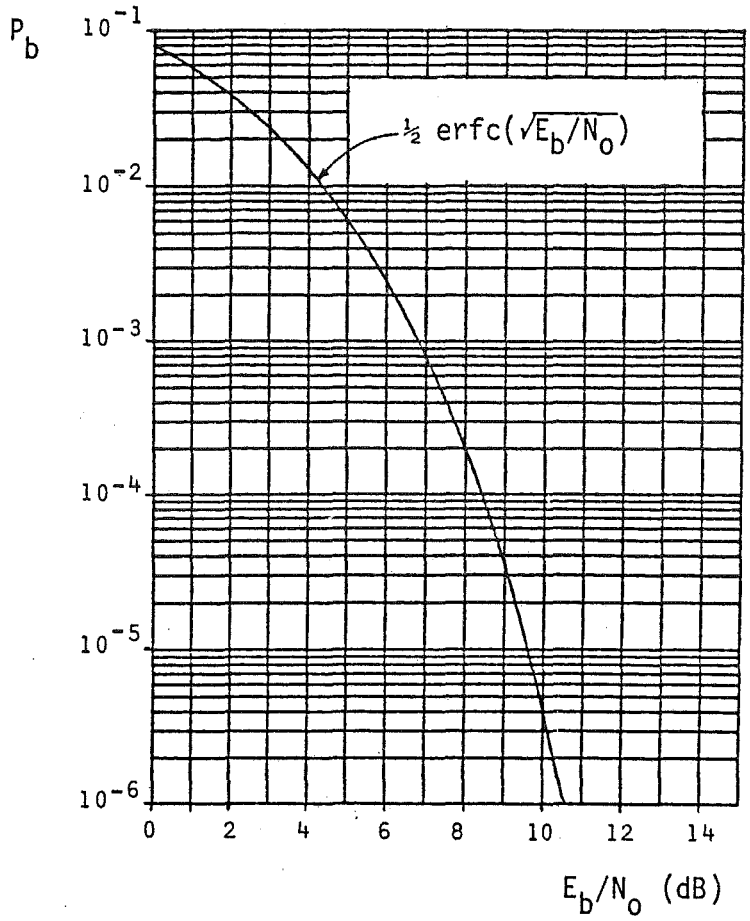


Fig. 2.8 BER curve for baseband antipodal signalling over a BSC.

podal transmission system in section 2.4 as a datum, a power efficient system could be considered as one with a BER performance curve to the left of that shown in Fig. 2.8, and a spectrally efficient system as one which requires less bandwidth per information bit.

2.5.1 Coding for Increased Power Efficiency

If a channel is power limited the encoder is used to add intentional redundancy to the information bits to effect a lower P_b for a given E_b/N_o. The redundancy is added in a time-wise manner resulting in a symbol-to-information bit ratio greater than 1. A reduction in spectral efficiency must therefore be accepted and the decoder at the transmitter requires an increase in complexity over the simple bit-by-bit decoder.

Since there are more symbols transmitted than information bits input, not all of the possible symbol combinations need to be used to convey the bit information. The reduction in P_b for a given E_b/N_o is achieved by making the sequence of symbols used to convey a given sequence of information bits as different as possible from any symbol sequences used for other bit combinations. There are two basic methods of encoding this redundancy into the symbols; namely, block codes and convolutional codes.

As the name suggests, block encoding entails coding D information bits into C binary symbols, with C > D. C symbols comprise a block which is independent of other symbol blocks. There are 2^C alternative blocks available, but only 2^D are required for coding. Clearly, the greater the difference or distance between the blocks used, the less likelihood of the decoder making a decision error in which block, hence which bits, were transmitted. Berlekamp [1980] and Pless [1982] give detailed discussion on some of the block codes available.

With convolutional encoding, the coded symbols are obtained by effectively convolving the information bits with a shift register stage [Heller, 1971; Ristenbatt, 1973; Viterbi, 1971]. One or more information bits are sequentially moved into a shift register, with the coded symbols taken from exclusive-OR operations on the shift register outputs. A simple example of this is shown in Fig. 2.9. With this example, 2 symbols are output each time 1 information bit is input. The decoder makes use of the fact that not all possible symbol combinations can occur (i.e., there is redundancy), and it has knowledge of these allowable alternatives as determined by the exclusive-OR connections on the shift register stage.

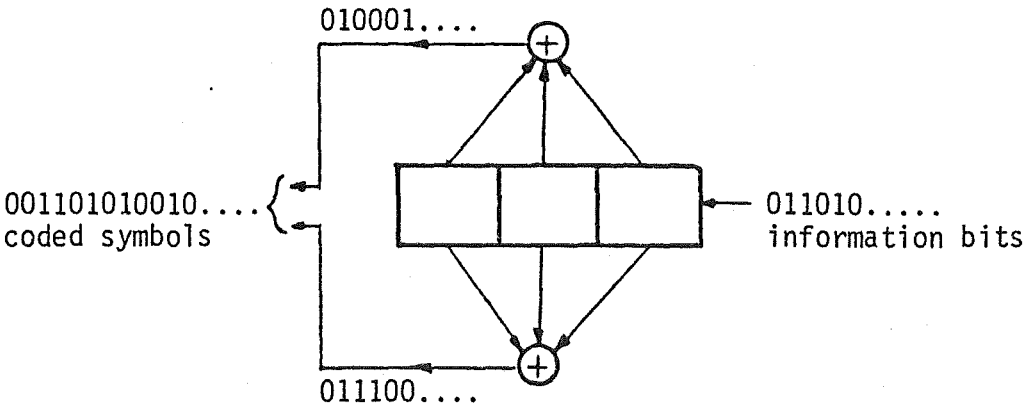


Fig. 2.9 Example of a convolutional coder. Logical levels "0" and "1" correspond to data bit levels -1 and +1, respectively.

2.5.2 Coding for Increased Spectral Efficiency

2.5.2.1 Basic multi-amplitude coding

With baseband systems the only method by which spectral efficiency can be improved is to increase the number of amplitude levels in the symbol alphabet beyond 2.

The simplest form of amplitude coding is where successive blocks of M information bits are each encoded into 1 of 2^M different symbols. An example of a 4-symbol alphabet is shown in Table 2.1. This simple coding method gives a 2-fold reduction in the channel signalling rate, hence a corresponding increase in spectral efficiency compared with 2-level signalling.

If the zero ISI criterion is adhered to, decoding can be carried out on a symbol-by-symbol basis similar to the antipodal signalling case. However, now the decoder must decide on which of the 2^M possible symbol levels each noise contaminated sample is nearest. The separation of the symbol levels will decrease if E_b is held fixed as M increases. The E_b/N_o degradation of a multi-amplitude system must therefore increase with M. The BER performance curve of an M-level system will move to the

right of the curve in Fig. 2.8 by the value of the E_b/N_o degradation. As an example, a 4-level system will suffer a 4 dB E_b/N_o degradation compared with antipodal signalling [Oetting, 1979].

TABLE 2.1

EXAMPLE OF AMPLITUDE CODING USING A 4-SYMBOL ALPHABET

Information Bits	Encoding Symbols
(-1, -1)	-3A
(-1, 1)	-A
(1, 1)	+A
(1, -1)	+3A

2.5.2.2 Partial response signalling (PRS)

PRS is often referred to under the alternative name of correlative level coding. The latter name more suitably describes the form of a PRS system which uses multi-amplitude coding where the coded symbols are correlated. This method relaxes the restraint of zero ISI by allowing ISI in a controlled manner and since this form of interference is known its effect can be removed by the decoder.

Lender [1963, 1964, 1966] introduced PRS as a data transmission method. Kretzmer [1966] categorised the characteristics of several PRS schemes and compared them with the ideal antipodal signalling system. Kabal and Pasupathy [1975] presented a unified study of PRS system, extending the analysis by Kretzmer. Initial studies into the BER performance of PRS systems were based on the precoding technique devised by Lender [1963] which enables bit-by-bit decoding without regard for the correlation among symbol levels. This form of decoding suffers an E_b/N_o degradation due to the multi-amplitude format of the signalling.

Kobayashi [Jan., Sept., Dec., 1971] recognised this and applied ML decoding to make full use of the information contained in the correlations, thus recovering a large portion of the E_b/N_o degradation the PRS system would otherwise suffer. Both precoding and ML decoding are discussed in more detail in Chapter 6 when they are applied to the QVSB system. An alternative to these two decoding methods is ambiguity zone detection which was introduced by Kobahashi and Tang [Aug., 1971]. This method makes use of the symbol correlations to reconstruct unreliable samples. If a sample contaminated with noise is in an "ambiguity zone" halfway between noiseless levels, it is rejected and the neighbouring samples are used to reconstruct the signal.

The encoding operation for PRS can be described by the system polynomial [Kabal and Pasupathy, 1975]

$$W(D) = \sum_{n=0}^{L-1} w_n D^n, \tag{2.60}$$

where D is the delay operator. More specifically, the encoding operation will have the form

$$c_\ell = w_{L-1} d_{\ell-L+1} + w_{L-2} d_{\ell-L+2} + \ldots + w_o d_\ell. \tag{2.61}$$

Equ. (2.61) shows c_ℓ to be a weighted sum of L information bits. Thus, c(t) is effectively obtained by convolving d(t) with a transversal filter which has L output taps weighted and summed according to (2.60). Once encoding is performed the symbols are transmitted without further ISI, so the system passband filter must be shaped to comply with Nyquist's first criterion for zero ISI. However, since convolution is associative, the transversal filter can be combined with the symbol shaping filter to produce a filter passband G(f) which incorporates the encoding operation. As a simple example consider a PRS system polynomial

$$W(D) = 1 + D, \tag{2.62}$$

which describes the duobinary, or class-1, PRS system. A transversal filter with transfer function given by (2.62) has spectral characteristics shown by the solid line in Fig. 2.10(a). Since, for this PRS system, the symbol rate equals the information bit rate, a rectangular passband shown by dashed lines in Fig. 2.10(a) would avoid any further ISI. Combining both the transversal and rectangular filters into the one filter shown in Fig. 2.10(b) gives a system passband G(f) that is more easily realised than the rectangular filter and incorporates the duobinary encoding. For optimum noise immunity, G(f) will be divided equally between transmitter and receiver. The time response g(t) for the duobinary system is shown in Fig. 2.10(c). The dashed lines in this figure illustrate the controlled ISI that would occur at the sampling points if bit sequence (...,-1,1,1,-1,1,...) was transmitted.

2.6 DIGITAL CARRIER-MODULATED SYSTEMS

In many instances the frequency spectrum available in the transmission channel does not match that of the baseband data. The baseband data system analysed in the previous sections can be modified to include modulation/demodulation stages. In carrier-modulated systems the amplitude, frequency and phase of a sinusoidal carrier are the parameters available for modification by the baseband data. Numerous modulation techniques using each of these parameters and their various combinations have been devised [Bennett and Davey, 1965; Clark, 1976; Haykin, 1978; Lucky et al., 1968].

The techniques available for increasing the power or spectral efficiency of baseband systems can also be applied directly to carrier-modulated systems. In this section a general model for carrier-modulated

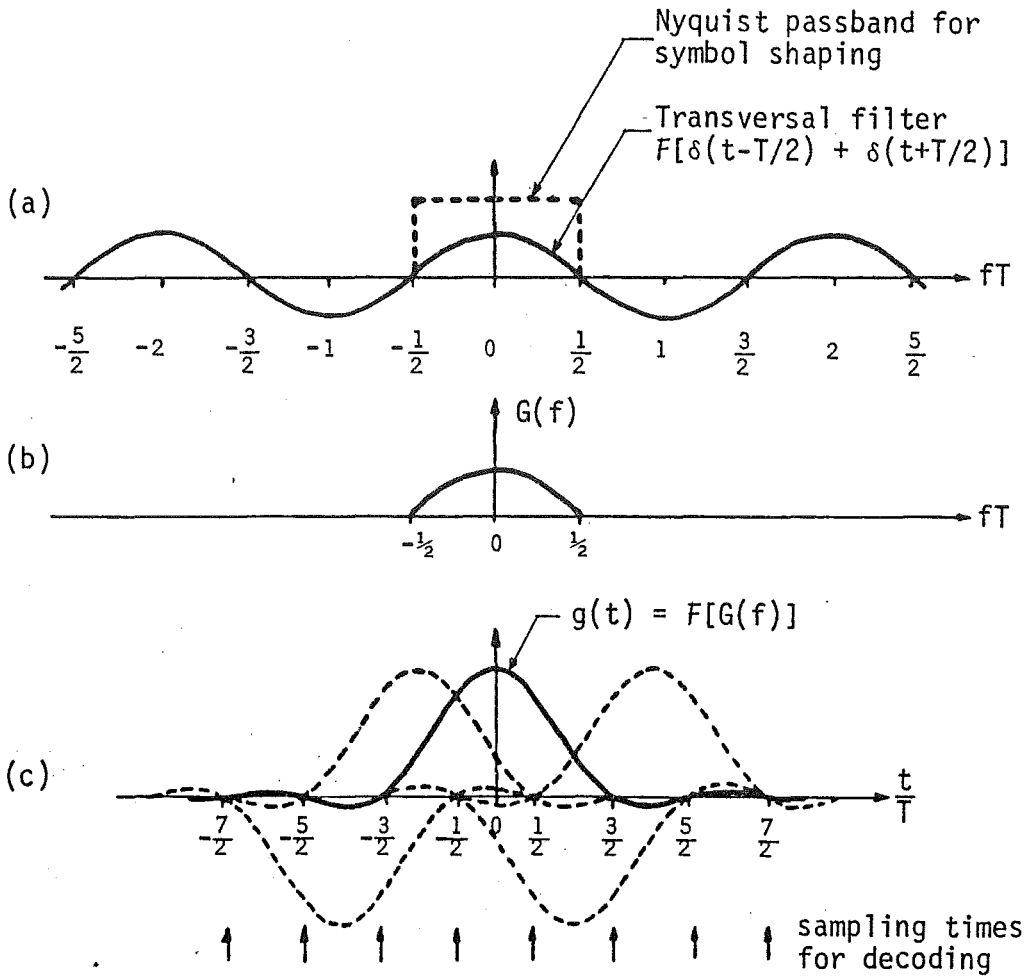


Fig. 2.10 Duobinary system signal shaping. (a) Separate transversal and symbol shaping filters which combine to give (b), and the resulting time response (c).

systems is used to introduce several techniques which, in addition to those used with baseband systems, increase spectral efficiency.

2.6.1 General Model

Carrier-modulated data systems can be described by the general model shown in Fig. 2.11. Considered here are linear systems employing coherent demodulation. By combining the quadrature related modulation paths all such modulation schemes can be treated using this model [Ungerboeck, 1974].

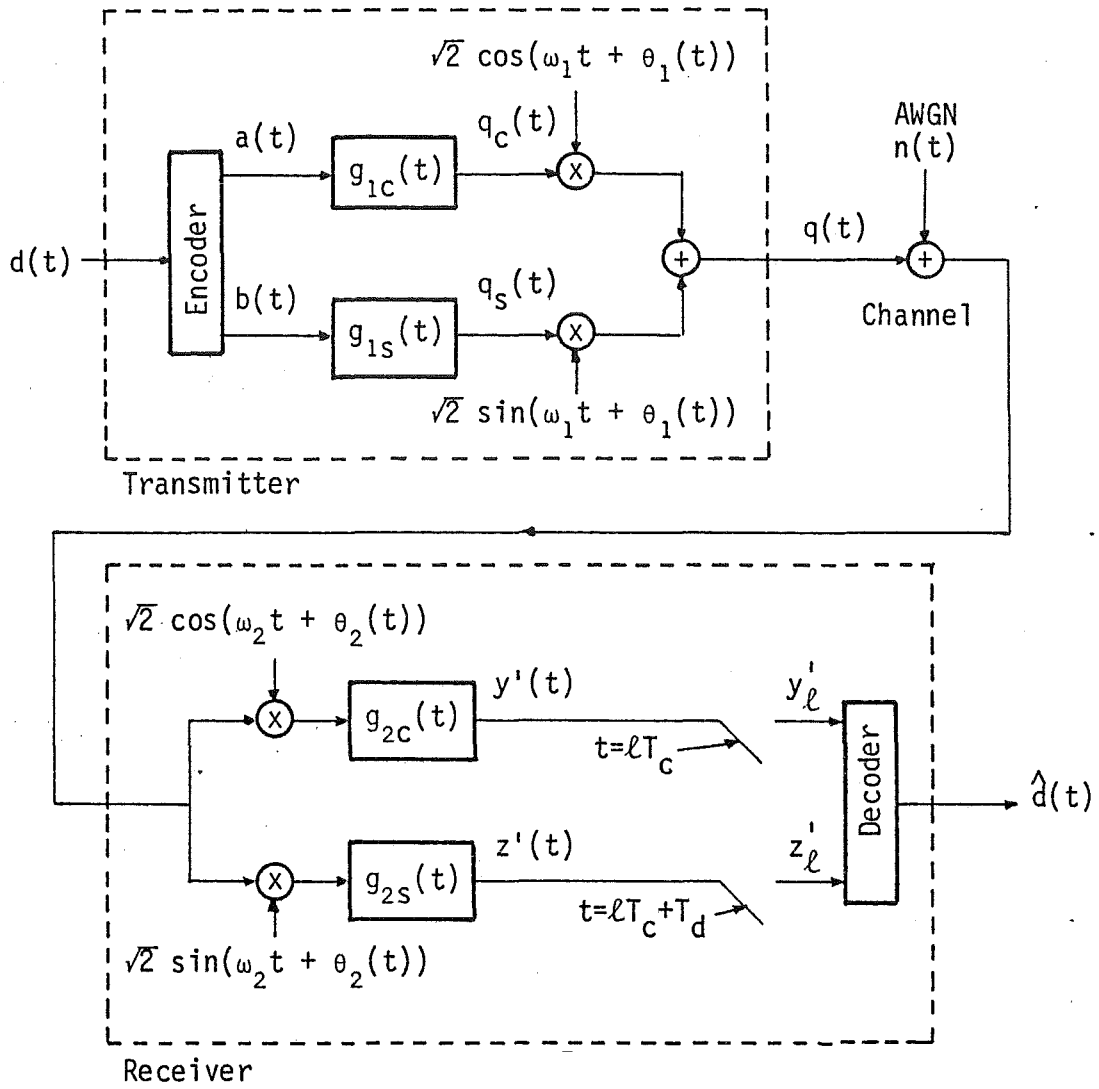


Fig. 2.11 General model for carrier-modulated data transmission systems.

In Fig. 2.11 $g_{1c}(t)$, $g_{1s}(t)$, $g_{2c}(t)$ and $g_{2s}(t)$ are all baseband filters which govern the symbol shaping. $y'(t)$ and $z'(t)$ are given by

$$y'(t) = y(t) + n_y(t) \tag{2.63}$$

and

$$z'(t) = z(t) + n_z(t), \tag{2.64}$$

where $y(t)$ and $z(t)$ are the desired signals, and $n_y(t)$ and $n_z(t)$ are the

interference due to the AWGN in the channel. The coded symbols are

$$a(t) = \sum_{\ell} a_{\ell}\delta(t - \ell T_c) \tag{2.65}$$

and

$$b(t) = \sum_{\ell} b_{\ell}\delta(t - \ell T_c - T_d), \tag{2.66}$$

where T_d represents any time offset between the quadrature paths. With most carrier-modulated systems the local oscillators at the transmitter and receiver are set to the same frequency, with $\omega_1 = \omega_2$. With coherent demodulation there is, ideally, no phase error so $\theta_1(t) = \theta_2(t)$. On this assumption there is no loss of generality in letting $\theta_1(t) = \theta_2(t) = 0$.

2.6.2 Binary Phase-Shift Keying (BPSK)

BPSK is the carrier system equivalent of the baseband antipodal signalling system given in section 2.4. It requires only a single modulation/demodulation path, with $b(t) = z(t) = 0$ and $a(t) = d(t)$. The channel signal is given by

$$q(t) = q_c(t)\ \sqrt{2}\ \cos(\omega_1 t)$$

$$= \{d(t) * g_{1c}(t)\}\ \sqrt{2}\ \cos(\omega_1 t). \tag{2.67}$$

After coherent demodulation by $\sqrt{2}\ \cos(\omega_1 t)$, the received signal is

$$y'(t) = d(t) * g_{1c}(t) * g_{2c}(t) + \{n(t)\ \sqrt{2}\ \cos(\omega_1 t)\} * g_{2c}(t). \tag{2.68}$$

Multiplying n(t) by $\sqrt{2}\cos(\omega_1 t)$ does not change the resulting AWGN spectral density from $N_0/2$, thus (2.68) describes identical filtering operations to those used with baseband systems. The same requirements for matched filtering, zero ISI and decoding therefore hold and the same BER performance is achieved (see Fig. 2.8).

BPSK, however, is a double sideband (DSB) signalling technique, so as shown by the example in Fig. 2.12 its spectral efficiency is halved compared with its equivalent baseband system.

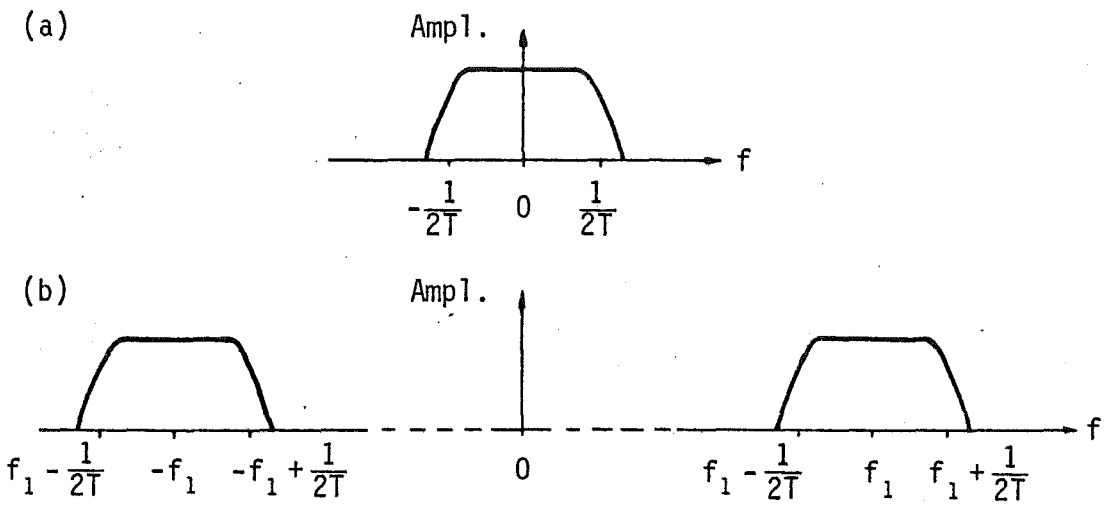


Fig. 2.12 Example of channel passbands required for (a) baseband antipodal signalling and (b) BPSK.

2.6.3 Quaternary Phase-Shift Keying (QPSK)

QPSK consists of two BPSK signals in phase quadrature, where

$$q(t) = \{a(t) * g_{1c}(t)\}\sqrt{2}\cos(\omega_1 t) + \{b(t) * g_{1s}(t)\}\sqrt{2}\sin(\omega_1 t). \tag{2.69}$$

Demodulating (2.69) by $\sqrt{2}\cos(\omega_1 t)$ and $\sqrt{2}\sin(\omega_1 t)$ gives, respectively,

$$y(t) = a(t) * g_{1c}(t) * g_{2c}(t) \tag{2.70}$$

and

$$z(t) = b(t) * g_{1s}(t) * g_{2s}(t) \tag{2.71}$$

In general, $g_{1c}(t)$ and $g_{1s}(t)$ are identical filters, which are in turn matched to, respectively, $g_{2c}(t)$ and $g_{2s}(t)$.

The quadrature related channels are independent of each other as long as there is no phase error in the demodulating carriers. Each quadrature related channel of the QPSK signal will therefore have identical BER performance to BPSK and hence to baseband antipodal signalling. For a given information rate, encoding d(t) for QPSK transmission involves dividing the information bits equally between a(t) and b(t). With the symbol rate in each path at half that for BPSK, a doubling of spectral efficiency is possible.

A particular form of QPSK is offset QPSK (OQPSK), where symbol sequences a(t) and b(t) are offset by $T_d = T$ seconds. The offset avoids nulls in the envelope of q(t) when narrow band filtering is used. As shown in Chapter 5, this is an advantage when nonlinear channel distortion is present.

2.6.4 Multi-Amplitude and Multi-Phase Signals

With baseband systems, increasing spectral efficiency is achieved by using a multi-amplitude symbol alphabet. The same method can also be used with carrier-modulated systems. Multi-amplitude signals can be used in both quadrature paths to produce quadrature-amplitude modulation (QAM). As for baseband systems, an increase in E_b/N_o degradation occurs as the number of symbol amplitudes is increased.

An increase in spectral efficiency can also be achieved by keeping the envelope of the modulated carrier fixed and encoding the data using the carrier phase. This is the form of M-ary PSK. In Fig. 2.11, q(t) can be written as

$$q(t) = A(t)\ \sqrt{2}\ \cos(\omega_1 t + \theta(t)), \tag{2.72a}$$

where

$$A(t) = \{q_c^2(t) + q_s^2(t)\}^{\frac{1}{2}} \tag{2.72b}$$

and

$$\theta(t) = -\tan^{-1}\left[\frac{q_s(t)}{q_c(t)}\right]. \tag{2.72c}$$

With M-ary PSK, A(t) is kept fixed and the symbol alphabet is derived from different values of $\theta(t)$. QPSK with rectangular symbol shaping is a particular form of M-ary PSK, where four phases 0, $\pm\pi/2$ and π radians make up the symbol alphabet.

The combined use of amplitude and phase-shift keying (APK) has been used to produce numerous signal encoding formats. Thomas et al. [1974] analyses the performance of several APK signalling schemes with large symbol alphabets. An excellent comparison and bibliography on various signalling schemes has been presented by Oetting [1979].

2.6.5 Single Sideband (SSB) and Vestigial Sideband (VSB) Signalling

With DSB systems each sideband contains identical information. A doubling in spectral efficiency can therefore be achieved by removal of one of the sidebands prior to signal transmission. The remaining sideband will contain sufficient information to enable the decoder to recover the transmitted data.

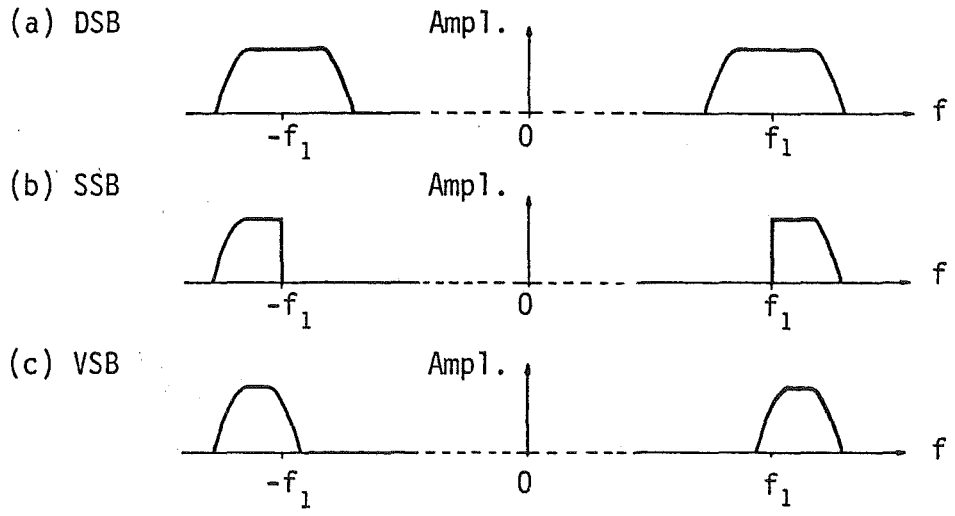


Fig. 2.13 Reducing (a) DSB spectrum to (b) SSB and (c) VSB.

SSB and VSB signal shaping are generally applied only to DSB data signals which have the form

$$q(t) = \{a(t) * g_{1c}(t)\} \sqrt{2} \cos(\omega_1 t). \qquad (2.73)$$

BPSK and multi-amplitude signalling, where no quadrature component is present, are included in this general form. They typically have a DSB spectrum as shown in Fig. 2.13(a) and unless special coding and/or filtering is used the DSB signal will have a significant amount of signal power at and around its carrier frequency f_1 Hz. Complete removal of one sideband while retaining the other for SSB signalling, as shown in Fig. 2.13(b), would therefore require unrealisable vertical rolloff filtering characteristics. This problem can be overcome, while retaining most of the spectral efficiency of SSB, by using VSB shaping. With VSB, one sideband plus a vestige of the other are transmitted as shown in Fig. 2.13(c). This enables gradual rolloff, hence realisable filters to be used in the signal shaping.

SSB and VSB signals can be defined as follows:

$$q(t) = \{a(t) * g_{1c}(t)\}\ \sqrt{2}\cos(\omega_1 t) \pm \{b(t) * g_{1s}(t)\}\ \sqrt{2}\sin(\omega_1 t), \tag{2.74a}$$

with

$$a(t) = b(t), \tag{2.74b}$$

where for SSB

$$g_{1s}(t) = H[g_{1c}(t)], \tag{2.75}$$

and for VSB

$$g_{1s}(t) = m(t) * H[g_{1c}(t)]. \tag{2.76}$$

In (2.74a) the plus sign applies to transmitting the lower sideband, whereas the minus sign applies to transmitting the upper sideband. $m(t)$ in (2.76) is used to modify the Hilbert transform function, hence the filtering characteristics from SSB to VSB. The VSB signal is analysed in detail in the following chapter.

Equ. (2.74a) indicates that SSB/VSB shaping introduces a quadrature component to the signal. This is usually regarded as an unwanted component which can only be removed with coherent demodulation by $\sqrt{2}\cos(\omega_1 t)$. SSB or VSB filtering of quadrature-modulated signals (e.g., QPSK) is generally avoided since the additional quadrature components introduced by the filtering would cause mutual interference between the quadrature related signals. However, this mutual interference is in a controlled form similar to PRS signalling and can, in theory, be removed by the decoder. A system which combines both quadrature modulation and VSB filtering is presented in Chapter 6.

CHAPTER 3

METHODS OF GENERATING AND DETECTING VSB SIGNALS

3.1 INTRODUCTION

The basic concepts of VSB data transmission are generally well understood. Several alternatives for transmitter and receiver configurations have appeared in literature, but in separate presentations. This chapter is intended as a concise reference to these alternatives, including some methods which have not specifically appeared in previous literature. Multi-level signalling is not considered here, since it is a simple extension of the antipodal case with performance degradations as for baseband signalling.

This study of various VSB techniques stems from attempts to design and construct a 9600 bits/sec duplex modem for operation over the subscriber telephone loop. Discussion of the modem design and difficulties encountered are given at the end of this chapter.

3.2 CONVENTIONAL VSB SYSTEM

The model for the conventional VSB data system is given in Fig. 3.1. Signal shaping can be shared by the lowpass filters $\ell(t)$ and bandpass filters $h_1(t)$ and $h_2(t)$, but in this analysis it will be assumed the bandpass filters perform the signal shaping and the presence of the lowpass filters is simply to remove unwanted high frequency terms.

With reference to Fig. 3.1, let

$$v'(t) = v(t) + \beta(t), \tag{3.1}$$

where $v(t)$ is the desired signal at the output of $h_2(t)$ and $\beta(t)$ is the noise. Likewise, let

$$x'(t) = x(t) + \alpha(t) \tag{3.2}$$

where $x(t)$ is the desired baseband signal at the receiver and $\alpha(t)$ is the noise.

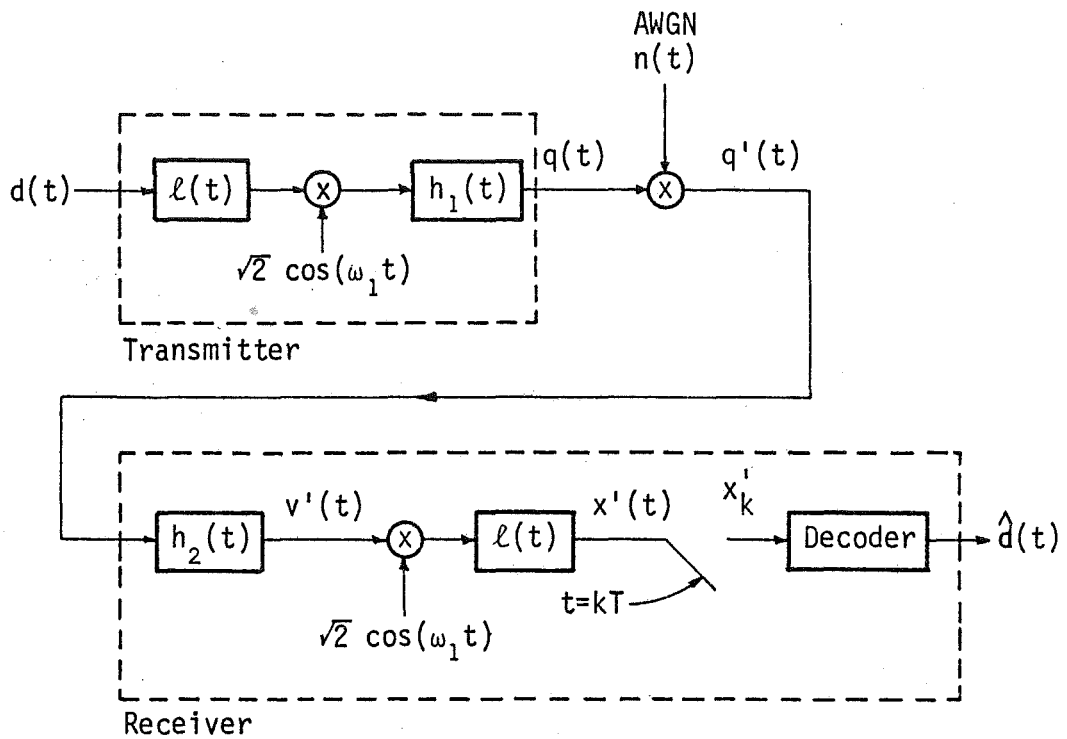


Fig. 3.1 Conventional VSB system.

3.2.1 Complex Representation

Since a VSB signal consists of both in-phase and quadrature components, it is convenient to make use of complex notation in its mathematical treatment.

The bandpass filters at the transmitter and receiver can be defined relative to the VSB carrier frequency as, respectively,

$$h_1(t) = h_{1c}(t)\cos(\omega_1 t) - h_{1s}(t)\sin(\omega_1 t)$$

$$= \mathrm{Re}\left[\{h_{1c}(t) + j\, h_{1s}(t)\}e^{j\omega_1 t}\right] \tag{3.3}$$

and

$$h_2(t) = h_{2c}(t)\cos(\omega_1 t) - h_{2s}(t)\sin(\omega_1 t)$$

$$= \mathrm{Re}\left[\{h_{2c}(t) + j\, h_{2s}(t)\}e^{j\omega_1 t}\right], \tag{3.4}$$

where the terms $h_{1c}(t)$, $h_{1s}(t)$, $h_{2c}(t)$ and $h_{2s}(t)$ are baseband functions which make up the complex envelopes of the filter responses. Letting

$$h(t) = h_1(t) * h_2(t), \tag{3.5}$$

then

$$h(t) = \frac{1}{2}\,\mathrm{Re}\Big[\{(h_{1c}(t) * h_{2c}(t) - h_{1s}(t) * h_{2s}(t))$$

$$+ j(h_{1c}(t) * h_{2s}(t) + h_{1s}(t) * h_{2c}(t))\}e^{j\omega_1 t}\Big] \tag{3.6a}$$

$$= \mathrm{Re}\left[\{h_c(t) + j\, h_s(t)\}e^{j\omega_1 t}\right], \text{ say.} \tag{3.6b}$$

The signal at the output of the transmitter filter is

$$q(t) = [\{d(t) * \ell(t)\}\sqrt{2}\cos(\omega_1 t)] * h_1(t)$$

$$= \frac{\sqrt{2}}{2}\,\mathrm{Re}\left[\{d(t) * (h_{1c}(t) + j\, h_{1s}(t))\}e^{j\omega_1 t}\right] \tag{3.7}$$

and the data signal at the output of the receiver filter is

$$v(t) = q(t) * h_2(t)$$

$$= \frac{\sqrt{2}}{2} \mathrm{Re}\left[\{d(t) * (h_c(t) + j\, h_s(t))\} e^{j\omega_1 t}\right]. \qquad (3.8)$$

After demodulating $v(t)$ by $\sqrt{2}\cos(\omega_1 t)$ the recovered baseband signal will be

$$x(t) = \{v(t)\ \sqrt{2}\cos(\omega_1 t)\} * \ell(t)$$

$$= \frac{1}{2} d(t) * h_c(t). \qquad (3.9)$$

Equ. (3.9) indicates that to achieve zero ISI, $h_c(t)$ must obey the same criterion as given in section 2.3.5 for baseband transmission.

Since the VSB system in Fig. 3.1 is linear, maximising the SNR of $x'(t)$ implies the SNR at the output of $h_2(t)$ must also be maximised. Match filtering therefore requires

$$h_2(t) = h_1(-t), \qquad (3.10)$$

corresponding to the frequency domain relationship

$$H_2(f) = H_1^*(f). \qquad (3.11)$$

Now, from (3.5) and (3.11)

$$H(f) = H_1(f)\ H_2(f)$$

$$= H_1(f)\ H_1^*(f)$$

$$= |H_1(f)|^2, \qquad (3.12)$$

so a requirement for maximising SNR at the receiver is that H(f) contains no imaginary part.

An example of how H(f) is made up of its various constituents is given in Fig. 3.2. In this case H(f) is a raised-cosine passband, so $H_1(f)$ and $H_2(f)$ will be identical cosine rolloff filters.

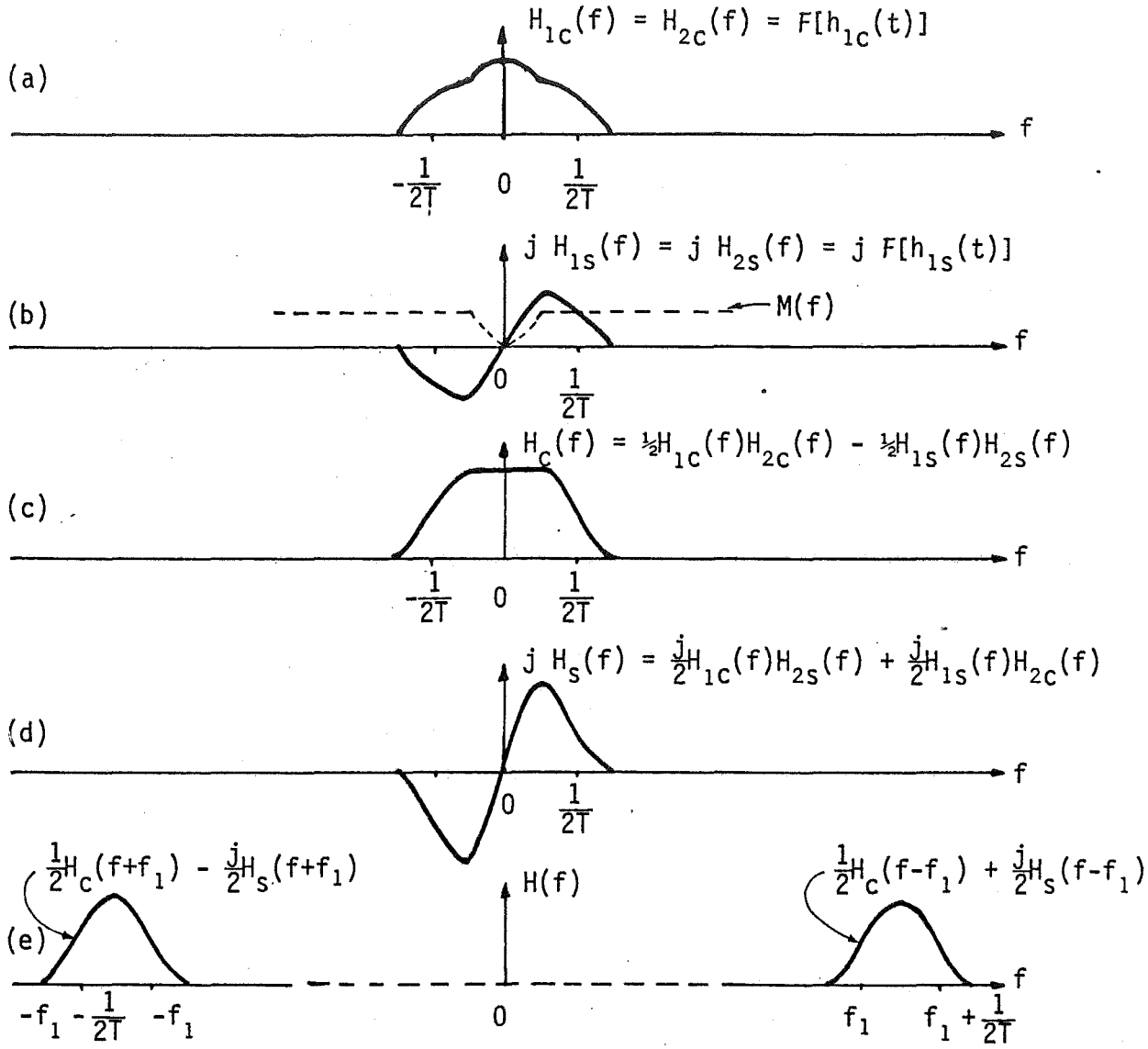


Fig. 3.2 Example of constituents which make up VSB passband H(f).

3.2.2 BER Performance

With zero ISI, each sample input to the decoder can be defined as

$$x_k' = x_k + \alpha_k, \tag{3.13}$$

where from (3.9)

$$x_k = d_k \frac{h_c(0)}{2}. \tag{3.14}$$

Letting

$$K = \frac{h_c(0)}{2}, \tag{3.15}$$

and σ_α^2 define the power of the noise term $\alpha(t)$, then the situation is identical to the baseband antipodal signalling case in section 2.4, with the probability of a bit error given by

$$P_b = \frac{1}{2} \operatorname{erfc}\left(\frac{K}{\sqrt{2}\ \sigma_\alpha}\right). \tag{3.16}$$

In Appendix 3 it is shown that for the VSB system

$$\sigma_\alpha^2 = \frac{N_o\ h_c(0)}{4}$$

$$= \frac{N_o}{2} K \tag{3.17}$$

and

$$E_b = \frac{h_c(0)}{2}$$

$$= K \tag{3.18}$$

which, on substituting into (3.14), gives

$$P_b = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_o}}\right). \tag{3.19}$$

The BER performance for VSB and baseband antipodal signalling are therefore identical.

3.2.3 Multiple Mixing Stages at the Transmitter

When the VSB carrier frequency is much larger than the signalling bandwidth there may be difficulties in realising practicable transmitter and receiver bandpass filters. The problem can be overcome for the transmitter filter by using two, or more, mixing stages [Haykin, 1978].

The two stage mixing procedure is illustrated in Fig. 3.3. The VSB signal is first generated at some low intermediate carrier frequency f_1', where realization of the bandpass shaping can be readily implemented. This signal is then DSB modulated so one of the sidebands is positioned at the desired VSB carrier frequency as shown in Fig. 3.3(b). The stopband between the two sidebands at the output of the second mixer enables the unwanted sideband to be removed with a practicable gradual rolloff filter.

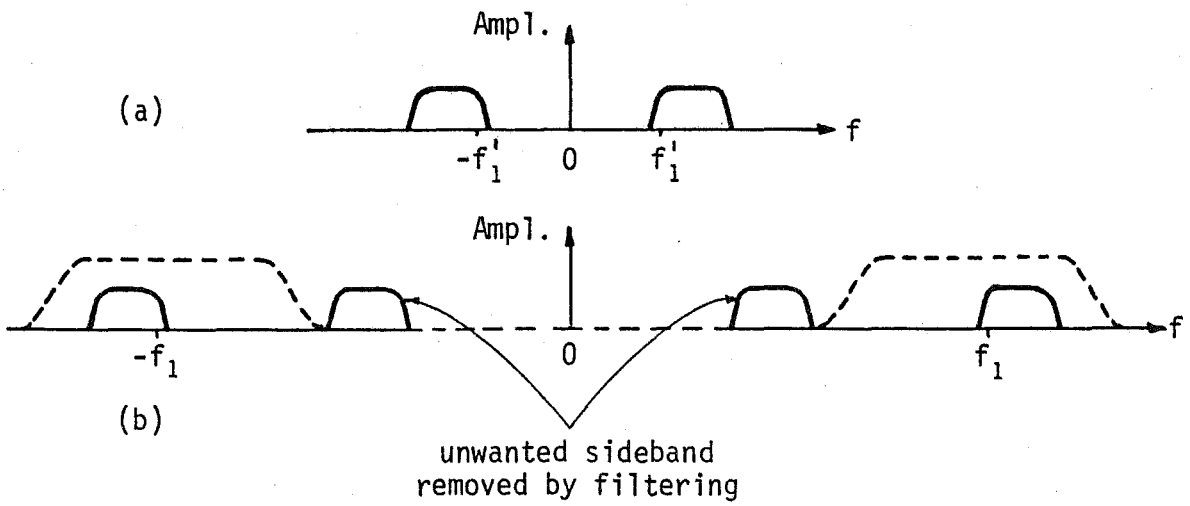


Fig. 3.3 Two mixing stages for generating a VSB signal. (a) Output of first stage where passband shaping is carried out. (b) Output of second stage with VSB passband at required carrier frequency and unwanted sideband removed.

3.3 COMPLEX MODULATION AND DEMODULATION

By complex modulation it is meant that the two quadrature related terms which make up the complex representation of VSB are generated separately and then combined to give the required signal. Similarly, complex demodulation refers to treating the quadrature related signals in the received VSB signal separately, then combining them prior to decoding. As will be shown, this enables VSB passband shaping to be carried out by baseband filters.

3.3.1 Complex Modulation

From (3.7),

$$q(t) = \left\{d(t) * \frac{h_{1c}(t)}{2}\right\} \sqrt{2}\cos(\omega_1 t) - \left\{d(t) * \frac{h_{1s}(t)}{2}\right\} \sqrt{2}\sin(\omega_1 t). \tag{3.20}$$

q(t) can therefore be generated using the stages shown in Fig. 3.4, so requiring no bandpass filtering at the mixer output.

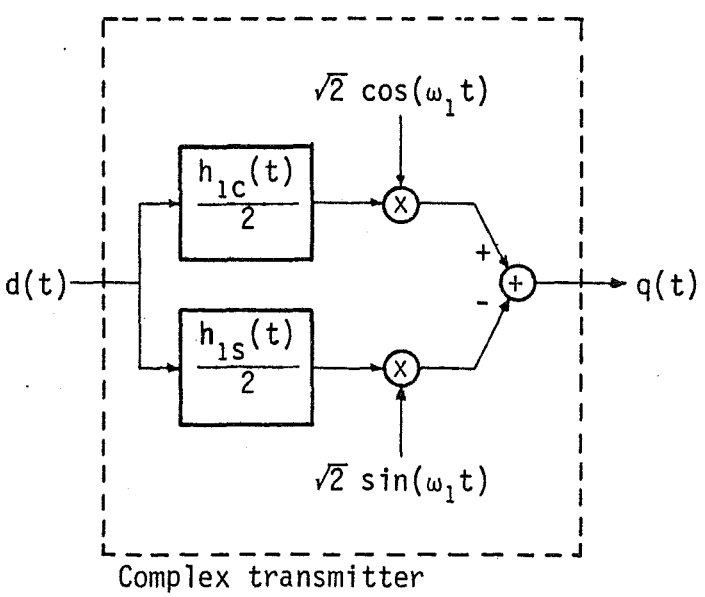


Fig. 3.4 Transmitter for VSB using complex modulation.

3.3.2 Complex Demodulation

The receiver can be treated in a similar manner to the transmitter. The baseband signal at the receiver is given by

$$x'(t) = [\{q'(t) * h_2(t)\}\ \sqrt{2}\ \cos(\omega_1 t)] * \ell(t), \tag{3.21}$$

where

$$q'(t) = q(t) + n(t). \tag{3.22}$$

Substituting the complex representation of $h_2(t)$ into (3.21) gives

$$\begin{aligned} x'(t) &= \left[\mathrm{Re}\left[q'(t) * \left\{(h_{2c}(t) + j\ h_{2s}(t))e^{j\omega_1 t}\right\}\right]\right. \\ &\qquad \left.\times\ \mathrm{Re}\left[\sqrt{2}\ e^{j\omega_1 t}\right]\right] * \ell(t) \\ &= \frac{\sqrt{2}}{2}\ \mathrm{Re}\left[\left\{q'(t)e^{-j\omega_1 t}\right\} * \left\{h_{2c}(t) + j\ h_{2s}(t)\right\}\right] \\ &= \{q'(t)\ \sqrt{2}\ \cos(\omega_1 t)\} * \frac{h_{2c}(t)}{2} \\ &\qquad + \{q'(t)\ \sqrt{2}\ \sin(\omega_1 t)\} * \frac{h_{2s}(t)}{2}\ . \end{aligned} \tag{3.23}$$

Equ. (3.23) describes the demodulation stages of the VSB receiver shown in Fig. 3.5. As for the complex modulator, all signal shaping is carried out by baseband filters.

3.3.3 Relationships Between the Complex Terms of the VSB Shaping Filters

The relationship between the two terms $h_{1c}(t)$ and $h_{1s}(t)$, which combine to make up the complex envelope of $h_1(t)$, can be described using the modified Hilbert transform:

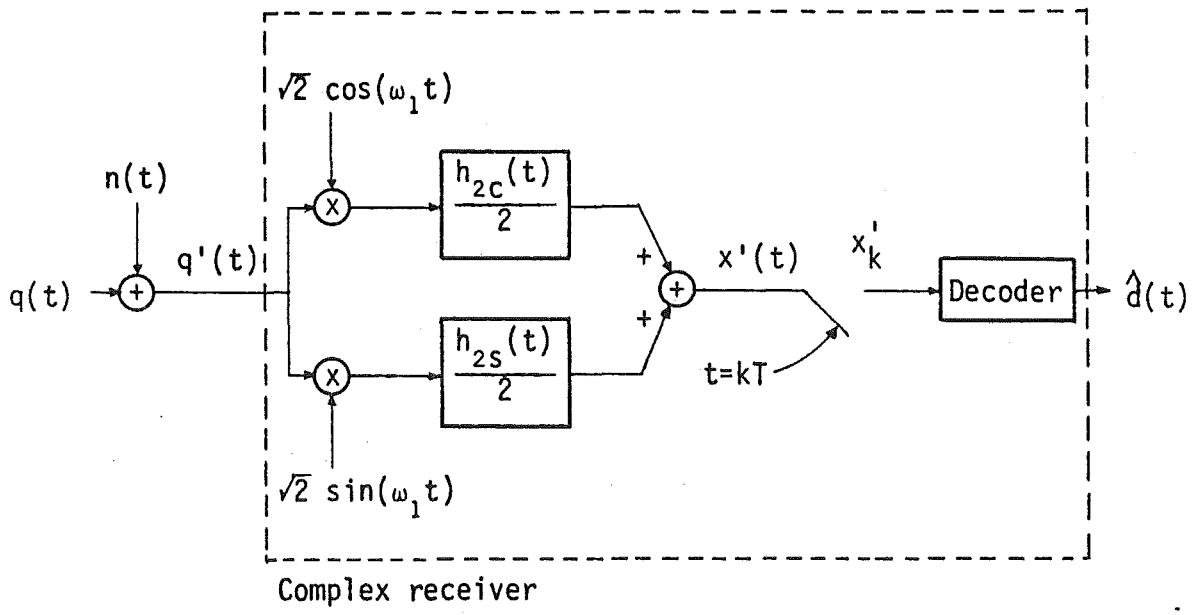


Fig. 3.5 Receiver for VSB using complex demodulation

$$h_{1s}(t) = m(t) * H[h_{1c}(t)], \tag{3.24}$$

or equivalently, in the frequency domain

$$H_{1s}(f) = -j\ \mathrm{sgn}(f)\ M(f)\ H_{1c}(f). \tag{3.25}$$

A similar relationship holds for $h_{2c}(t)$ and $h_{2s}(t)$, hence also for $h_c(t)$ and $h_s(t)$.

The function m(t) is used to modify the Hilbert transform of $h_{1c}(t)$ in a range of frequencies around f = 0 Hz. This is illustrated in Fig. 3.2, where $H_{1c}(f)$ and $H_{1s}(f)$ are related by (3.25), with M(f) shown by the dashed line in Fig. 3.2(b). A similar relationship is seen to exist between $H_c(f)$ and $H_s(f)$.

Hill [1975] used the modified Hilbert transform representation to arrive at several optimum pulse shapes for VSB signalling. Actual implementation of a VSB signal generator using the components of the complex envelope was considered by Fang [1972], where he showed that m(t) can be

used to modify amplitude or phase characteristics of the Hilbert transform to obtain the VSB passband shaping.

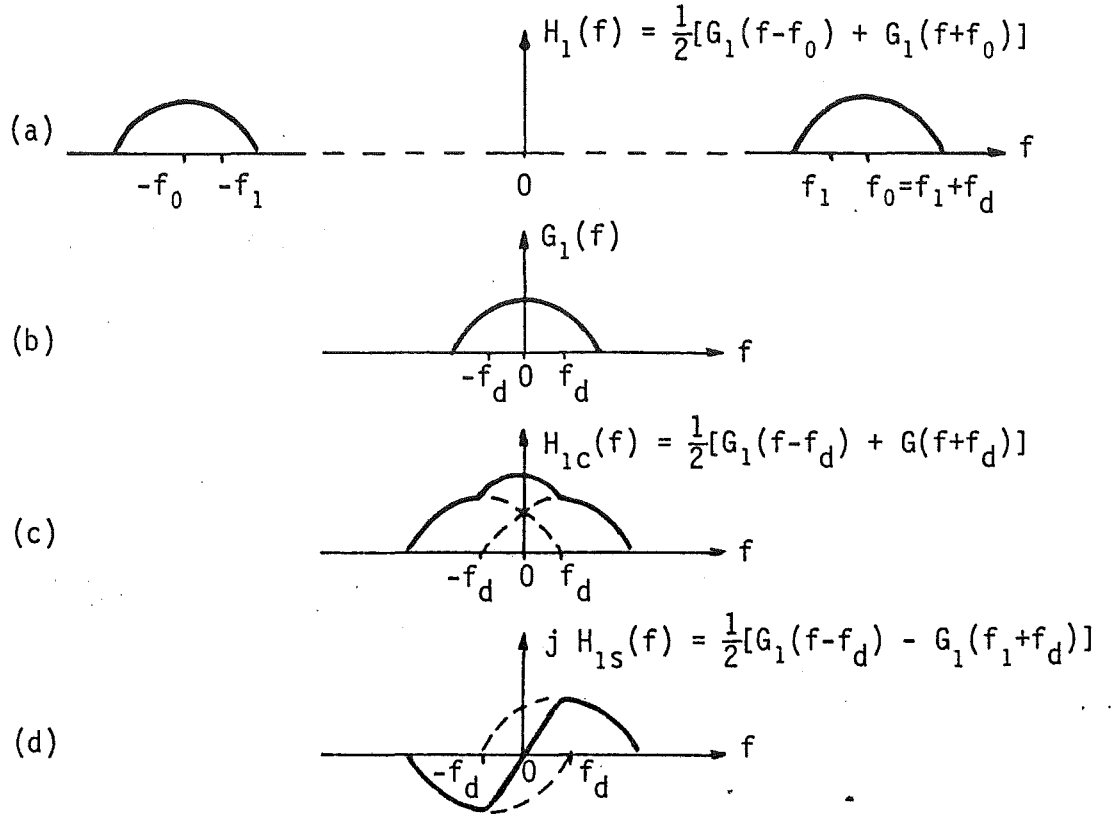


Fig. 3.6 Constituents of VSB passband when it is symmetrical about its center frequency.

A simple relationship exists between $h_{1c}(t)$ and $h_{1s}(t)$ if $h_1(t)$ defines a bandpass function which is symmetrical about its center frequency f_0. This is illustrated in Fig. 3.6, where

$$h_1(t) = g_1(t)\cos(\omega_0 t), \tag{3.26}$$

with $g_1(t)$ as a lowpass function. If the difference between the carrier and passband center frequencies is

$$\omega_d = \omega_0 - \omega_1, \tag{3.27}$$

then

$$h_1(t) = g_1(t)\cos(\omega_d t)\cos(\omega_1 t) - g_1(t)\sin(\omega_d t)\sin(\omega_1 t). \quad (3.28)$$

Comparing (3.3) and (3.28) gives the complex envelope terms as

$$h_{1c}(t) = g_1(t)\cos(\omega_d t) \quad (3.29a)$$

and

$$h_{1s}(t) = g_1(t)\sin(\omega_d t). \quad (3.29b)$$

Similar relationships will also hold for $h_2(t)$ and $h(t)$, where

$$h_2(t) = g_2(t)\cos(\omega_0 t) \quad (3.30)$$

and

$$h(t) = g(t)\cos(\omega_0 t), \quad (3.31)$$

with

$$g(t) = \frac{1}{2} g_1(t) * g_2(t). \quad (3.32)$$

3.4 IN-BAND GENERATION

In-band generation of data signals eliminates the need for separate baseband filtering, mixing and bandpass filtering stages at the transmitter. VSB is just one of a number of signalling formats which can be generated using in-band techniques [Choquet and Nussbaumer, 1971; Croisier and Pierret, 1970; Kalet and Weinstein, 1973; Weinstein, 1976]. Sunde [July, 1954] mentioned the possibility of generating symmetrical and asymmetrical sideband AM systems by exciting bandpass characteristics directly with a PAM signal, but he did not study this approach in depth,

considering it to pose too great a difficulty in implementation compared with conventional modulation techniques. However, the availability of digital processing techniques using large scale integrated circuitry now enables in-band generation to be a feasible alternative.

3.4.1 Restricted Carrier Frequency

Fig. 3.7 shows, in its simplest form, the in-band generation of VSB. The output signal is given by

$$\begin{aligned} q(t) &= d(t) * h_1(t) \\ &= \sum_{k=0}^{N} d_k \delta(t - kT) * \{h_{1c}(t) \cos(\omega_1 t) - h_{1s}(t) \sin(\omega_1 t)\} \\ &= \sum_{k=0}^{N} d_k \{h_{1c}(t - kT) \cos(\omega_1 (t - kT)) \\ &\quad - h_{1s}(t - kT) \sin(\omega_1 (t - kT))\}. \end{aligned} \tag{3.33}$$

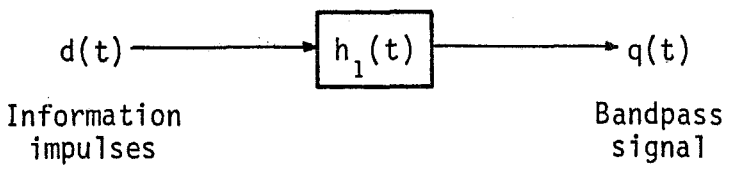


Fig. 3.7 In-band generation model.

Equ. (3.33) indicates the phase of the carrier is dependent on the timing of the data impulses. However, if the restriction is made that

$$f_1 T = M, \quad M \text{ an integer}, \tag{3.34}$$

the term $(\omega_1(t - kT))$ in (3.33) becomes

$$(\omega_1(t - kT)) = 2\pi f_1 t - 2\pi kM$$

$$= (2\pi f_1 t)_{\text{mod } 2\pi}, \qquad (3.35)$$

and (3.33) reduces to

$$q(t) = \sum_{k=0}^{N} \{d_k\, h_{1c}(t - kT)\} \cos(\omega_1 t) - \{d_k\, h_{1s}(t - kT)\} \sin(\omega_1 t). \qquad (3.36)$$

The carrier phase in (3.36) is now independent of the data impulses, with the restriction given by (3.34) ensuring continuity of the phase as each data pulse excites $h_1(t)$. Equ. (3.36) is identical, except for a constant, to (3.7), the equation for a VSB signal. With the restriction given by (3.34), the passbands available for in-band generation of VSB are shown in Fig. 3.8.

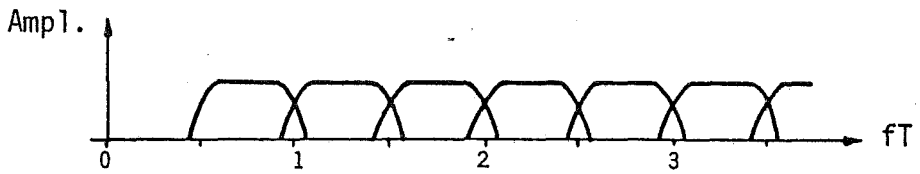


Fig. 3.8 Passbands available for in-band generation of VSB. f_1 must be some integer multiple of 1/T.

Croisier and Pierret [1970] put in-band generation into practice with a technique they called digital echo modulation. With this method $h_1(t)$ is defined by discrete samples. To avoid aliasing, the sampling theorem [Haykin, 1978; Wozencraft and Jacobs, 1965] generally requires the sampling frequency to be at least twice that of the highest significant frequency component of $h_1(t)$. This restricts practicable digital echo modulation systems to passbands where f_1T is small so their realisation does not require an excessive number of samples. A method to relax this

restriction is given in section 3.4.3.

There are two alternatives for the type of receiver which can be used when $h_1(t)$ is one of the restricted passbands shown in Fig. 3.8. The signal input to the receiver has a VSB format, so can be detected as such by mixing it down to baseband. The other possibility, which is employed with digital echo modulation, is to sample and decode the received signal at the output of a receiver bandpass filter $h_2(t)$. If the VSB passband has, for example, raised-cosine shaping as shown in Fig. 3.9(a), it obeys the critereon for zero ISI. Fig. 3.9(c) illustrates how overlapping responses due to d(t) do not interfere at the sampling points. Clearly, the tolerance in sampling error will decrease as f_1T increases.

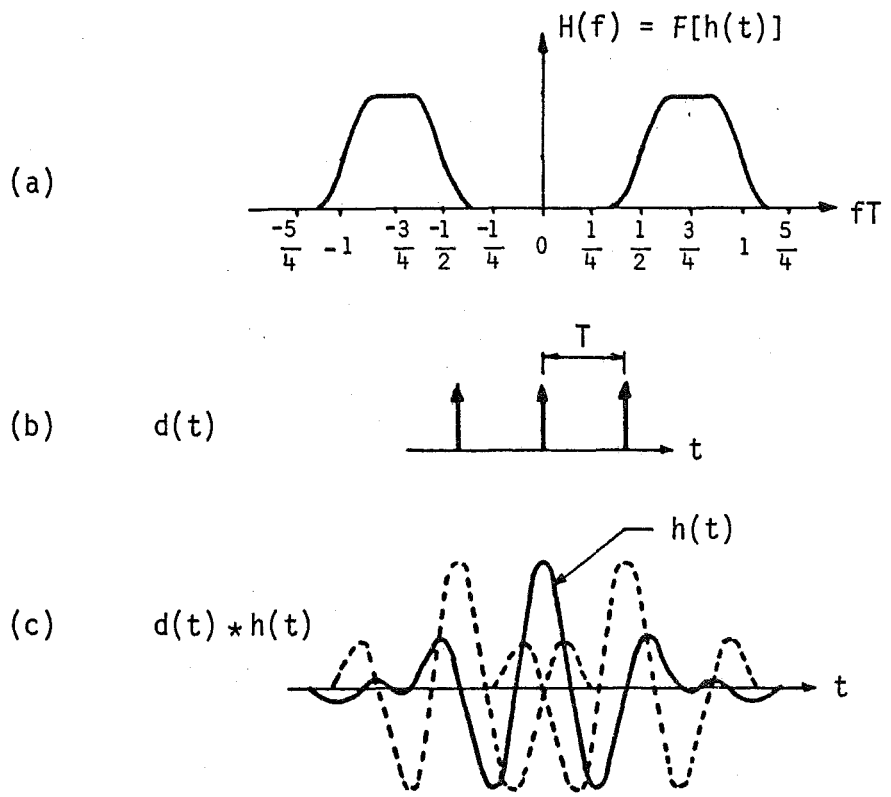


Fig. 3.9 (a) Raised-cosine passband for in-band generation. (b) Information bits. (c) Overlapping responses with points of zero ISI at sampling times.

If the required carrier frequency is much larger than, and

possibly not a multiple of 1/T, Croisier and Pierret [1970] use digital echo modulation for initial passband shaping, then employ a mixing stage at the transmitter to translate the signal up to the desired passband. With the in-band generated signal having no signal power about d.c., the unwanted sideband can be readily removed after the mixing operation. This is similar to the two stage mixing operation illustrated in Fig. 3.3, with the digital echo modulation stage replacing the first mixing stage which derives the intermediate VSB signal.

3.4.2 Unrestricted Carrier Frequency

If f_1 is not some integer multiple of 1/T, Choquet and Nussbaumer [1971] and Kalet and Weinstein [1973] showed in-band generation can still be used to derive the desired bandpass signal without need for a mixing stage. When

$$f_1 T = \frac{m_1}{m_2}, \tag{3.37}$$

with m_1 and m_2 as relative primes, (3.33) will not reduce to (3.36). The problem of ensuring carrier phase continuity as each information pulse excites $h_1(t)$ can be overcome by employing a bank of filters at the transmitter as shown in Fig. 3.10.

Defining the k^{th} filter in Fig. 3.10 as

$$h_1^{(k)}(t) = \mathrm{Re}\left[\{h_{1c}(t) + j\, h_{1s}(t)\} e^{j\omega_1 t}\, e^{j\phi_k}\right], \tag{3.38}$$

where ϕ_k is the phase adjustment, then the output of the filter will be

$$q^{(k)}(t) = d_k\{h_{1c}(t - kT)\cos(\omega_1 t - \omega_1 kT + \phi_k) - h_{1s}(t - kT)\sin(\omega_1 t - \omega_1 kT + \phi_k)\}. \tag{3.39}$$

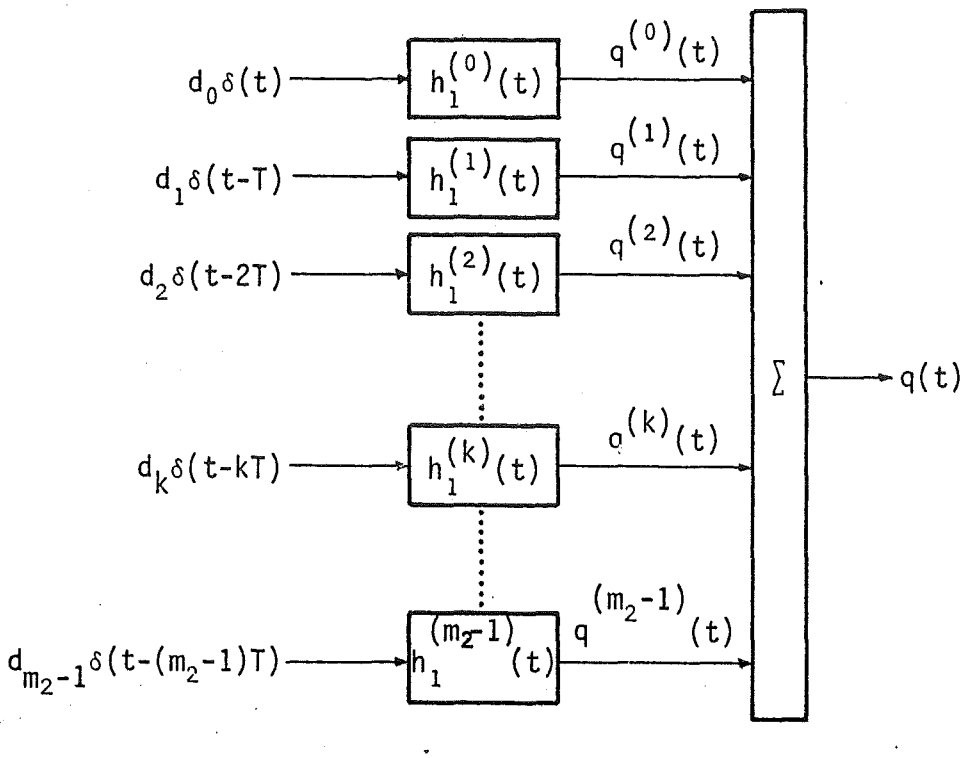


Fig. 3.10 In-band generation model when $f_1T = m_1/m_2$, with m_1 and m_2 relative primes.

If the carrier phase of $q(t)$ is to be independent of the bit timing the phase of $q^{(0)}(t)$ must equal that of $q^{(k)}(t)$; i.e.,

$$\omega_1 t + \phi_0 = \omega_1 t - \omega_1 kT + \phi_k, \tag{3.40a}$$

or

$$\phi_k = \omega_1 kT + \phi_0. \tag{3.40b}$$

Taking ϕ_0 as the reference phase and equal to zero gives

$$\phi_k = 2\pi f_1 kT$$

$$= 2\pi k \frac{m_1}{m_2}. \tag{3.41}$$

Since the modulo-2π value of ϕ_k is of interest, the phase adjustment of the k^{th} filter is

$$\phi_k = \left(2\pi k \frac{m_1}{m_2}\right)_{\text{mod } 2\pi}. \tag{3.42}$$

ϕ_k can take on m_2 distinct values, so $h^{(k)}(t)$ and $h^{(k+m_2)}(t)$ will be identical filters. In-band generation of $q(t)$ can therefore be achieved by cyclically exciting a bank of m_2 filters. Weinstein [1976] showed $q(t)$ could be generated using a single transversal filter with time-varying tap weightings.

3.4.3 Defining the Passband Using Discrete Samples

Implementation of in-band generation using digital techniques or transversal filters requires $h_1(t)$ and possibly $h_2(t)$ to be defined by discrete samples.

If a signal to be sampled has its highest significant frequency component at f_h Hz, then to avoid aliasing the sampling theorem in its most commonly stated form requires f_s, the sampling frequency, to obey the relationship

$$f_s \geq 2f_h. \tag{3.43}$$

Equ. (3.43) applies mainly to signals with lowpass frequency characteristics. It has been shown [Haykin, 1978; Panter, 1965] that for signals with a bandpass spectrum the minimum sampling frequency can be reduced to below that indicated by (3.43) and aliasing still avoided. For a general bandpass signal as shown in Fig. 3.11(a), where the highest significant signal frequency is f_h and the bandwidth is B, the minimum sampling frequency which avoids aliasing is given by the heavy line in Fig. 3.11(b).

An important point, which was illustrated by Kelly [1979], is that f_s may not take on all the values above this minimum sampling frequency; rather, f_s is restricted to frequencies indicated by the shaded regions in Fig. 3.11(b).

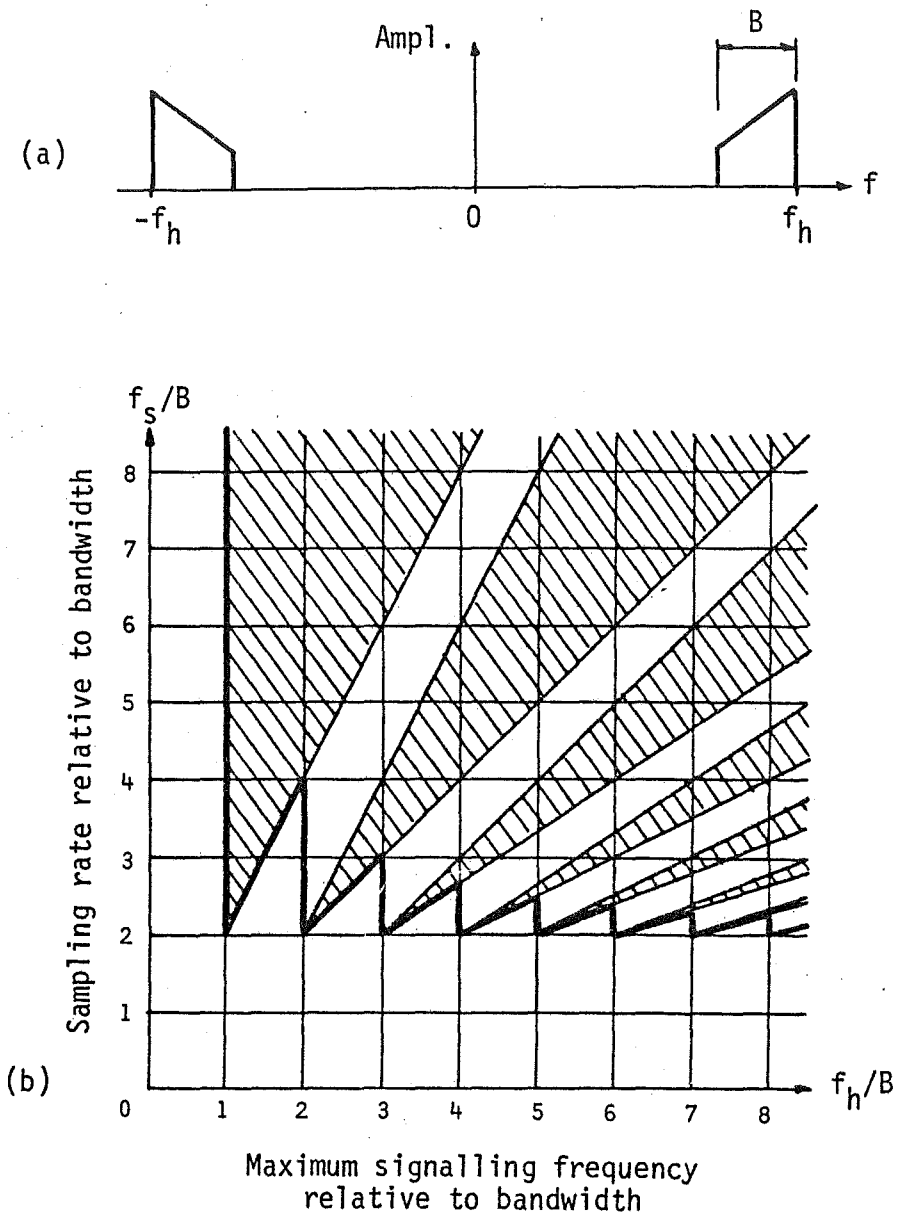


Fig. 3.11 (a) General bandpass spectrum. (b) Heavy line indicates minimum sampling frequency for a bandpass signal, and the shaded regions indicate allowable sampling frequencies.

3.4.4 Combined In-Band Generation and Complex Modulation

The use of a mixing stage following in-band generation was suggested as one method of deriving the desired signal when $f_1 >> 1/T$. Employing a single mixing stage after the in-band generator requires the inclusion of a post-mixer bandpass filter to remove the unwanted sideband. This can be dispensed with if complex modulation is used in a similar manner to the phase discrimination method for SSB.

The technique combining in-band generation and complex modulation is illustrated in Fig. 3.12(a). Intermediate signals $q_c'(t)$ and $q_s'(t)$, where

$$q_s'(t) = H[q_c'(t)], \tag{3.44}$$

are derived using in-band generation, resulting in their relative frequency spectra shown in Fig. 3.12(b) and Fig. 3.12(c). Following the mixing stages, $q_c''(t)$ and $q_s''(t)$ have relative frequency spectra shown in Fig. 3.12(d) and Fig. 3.12(e), respectively. Summation of these two DSB signals results in mutual cancellation of one of the sidebands leaving the desired VSB signal (Fig. 3.12(f)).

In Fig. 3.12(a), $h_{1c}'(t)$ and $h_{1s}'(t)$ are bandpass functions which govern the spectral shaping of $q_c(t)$ and $q_s(t)$, respectively. Implementation of this method requires good cancellation of the unwanted sideband, hence $h_{1c}'(t)$ and $h_{1s}'(t)$ must be closely matched as a Hilbert pair. Del Re [1981] gave a method of achieving this if the filters are defined by discrete samples. If $h_{1c}'(t)$ has a spectrum which is symmetrical about its center frequency f_0' Hz, then it will have the form

$$h_{1c}'(t) = g_1'(t) \cos(\omega_0' t), \tag{3.45}$$

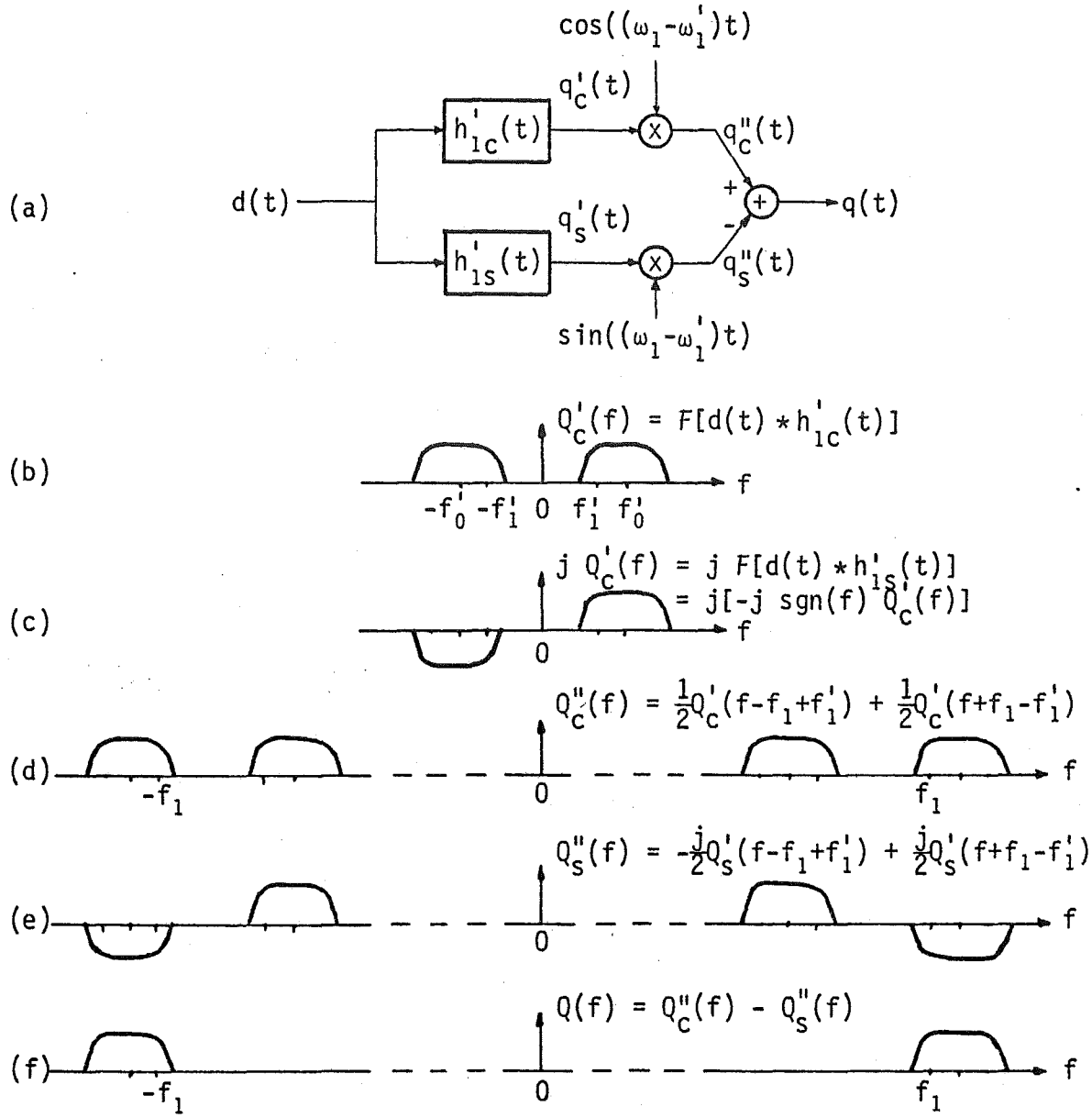


Fig. 3.12 (a) Model for combined in-band and complex generation of signal passband. (b) Through (f) illustrate spectral management of signals.

where $g'_1(t)$ is a lowpass function. $h'_{1s}(t)$ is therefore given by

$$h'_{1s}(t) = H[h'_{1c}(t)]$$

$$= g'_1(t)\,\sin(\omega'_0 t). \tag{3.46}$$

$g'_1(t)$ is common to both $h'_{1c}(t)$ and $h'_{1s}(t)$. Also, the two terms $\cos(\omega'_0 t)$

and $\sin(\omega_0' t)$ have the same absolute value when their arguments equal odd multiples of $\pi/4$ radians; or more specifically

$$\cos\left(\frac{n\pi}{4}\right) = (-1)^{\frac{|n-1|}{2}} \sin\left(\frac{n\pi}{4}\right), \quad n = \pm 1, \pm 3, \pm 5, \ldots \tag{3.47}$$

The multiples of $\pi/4$ radians correspond to times

$$t = \frac{n\pi}{4\omega_0'}$$

$$= \frac{n}{8f_0'}, \quad n = \pm 1, \pm 3, \pm 5, \ldots \tag{3.48}$$

At these points, $h_{1c}(t)$ and $h_{1s}(t)$ will obey the relationship:

$$h_{1s}'\left(\frac{n}{8f_0'}\right) = (-1)^{\frac{|n-1|}{2}} h_{1c}'\left(\frac{n}{8f_0'}\right), \quad n = \pm 1, \pm 3, \pm 5, \ldots \tag{3.49}$$

The implication of (3.49) is, if $h_{1c}(t)$ and $h_{1s}(t)$ are defined by samples spaced at intervals

$$T_s = 1/(4f_0') \text{ sec.}, \tag{3.50}$$

the same sample weightings can be used to define both filter responses, with relative polarity reversals of every second sample weighting as per (3.49). The actual sampling points relative to the nominal time origin of the filter responses are given by (3.48). $h_{1c}(t)$ and $h_{1s}(t)$ can therefore be closely matched by using the same weighting terms or network for both filters. Any errors in the weightings will be reflected in both $h_{1c}(t)$ and $h_{1s}(t)$, so preserving matching and ensuring good cancellation of the unwanted sideband.

3.5 GENERATION AND DETECTION OF VSB SIGNALS USING OQPSK TECHNIQUES

The similarity between VSB and quadrature modulation systems is evident from the complex modulation and demodulation models for VSB derived in section 3.3. van den Elzen [1975] showed that a VSB signal could be demodulated using OQPSK techniques. Clearly, if this is possible, a OQPSK modulator could be used to generate a VSB signal. Configurations for both a VSB transmitter and receiver employing OQPSK stages are derived here.

In these derivations the VSB passband will be assumed to be symmetrical about its center frequency f_0 Hz, where

$$h_1(t) = g_1(t) \cos(\omega_0 t) \tag{3.51}$$

and

$$h_2(t) = g_2(t) \cos(\omega_0 t), \tag{3.52}$$

with $g_1(t)$ and $g_2(t)$ as lowpass functions. The difference between f_0 and the VSB carrier f_1 will be defined as

$$f_d = f_0 - f_1 \text{ Hz}. \tag{3.53}$$

With optimum narrow band filtering

$$f_d = \pm \frac{1}{4T} \text{ Hz}. \tag{3.54}$$

3.5.1 Generation of VSB

Using (3.7) and (3.51), the VSB signal can be defined relative to its carrier frequency as

$$q(t) = \mathrm{Re}\left[\{d(t) * \ell(t)\}\ \sqrt{2}\ e^{j\omega_1 t}\right] * \mathrm{Re}\left[g_1(t)\ e^{j\omega_0 t}\right]$$

$$= \frac{\sqrt{2}}{2}\ \mathrm{Re}\left[\left\{d(t) * \left(g_1(t)\ e^{j\omega_d t}\right)\right\} e^{j\omega_1 t}\right]. \qquad (3.55)$$

Alternatively, q(t) can be defined relative to the VSB passband center frequency as

$$q(t) = \frac{\sqrt{2}}{2}\ \mathrm{Re}\left[\left\{d(t) * \left(g_1(t)\ e^{j\omega_d t}\right)\right\} e^{-j\omega_d t}\ e^{j\omega_0 t}\right]$$

$$= \frac{\sqrt{2}}{2}\ \mathrm{Re}\left[\left\{\left(d(t)\ e^{-j\omega_d t}\right) * g_1(t)\right\} e^{j\omega_0 t}\right]$$

$$= \left[\{d(t)\ \cos(\omega_d t)\} * \frac{g_1(t)}{2}\right] \sqrt{2}\ \cos(\omega_0 t)$$

$$+ \left[\{d(t)\ \sin(\omega_d t)\} * \frac{g_1(t)}{2}\right] \sqrt{2}\ \sin(\omega_0 t) \qquad (3.56)$$

Equ. (3.56) describes the model shown in Fig. 3.13(a). This is the general form of a QPSK signal generator with the inclusion of an encoding stage. As shown in Fig. 3.13(b), the encoding operation can be simplified by aligning the zero crossings of $\cos(\omega_d t)$ and $\sin(\omega_d t)$ with the occurrence of the data bits. The result is two bit streams, a(t) and b(t), at half the information rate and offset relative to each other by T seconds. This is the form of an OQPSK signal generator.

It is of interest to note this method of VSB generation is similar to that developed by Weaver [1956] as a third method of generating SSB.

3.5.2 Detection of VSB

If q'(t) is the VSB signal plus noise at the input to the receiver filter, the demodulated signal prior to sampling will be

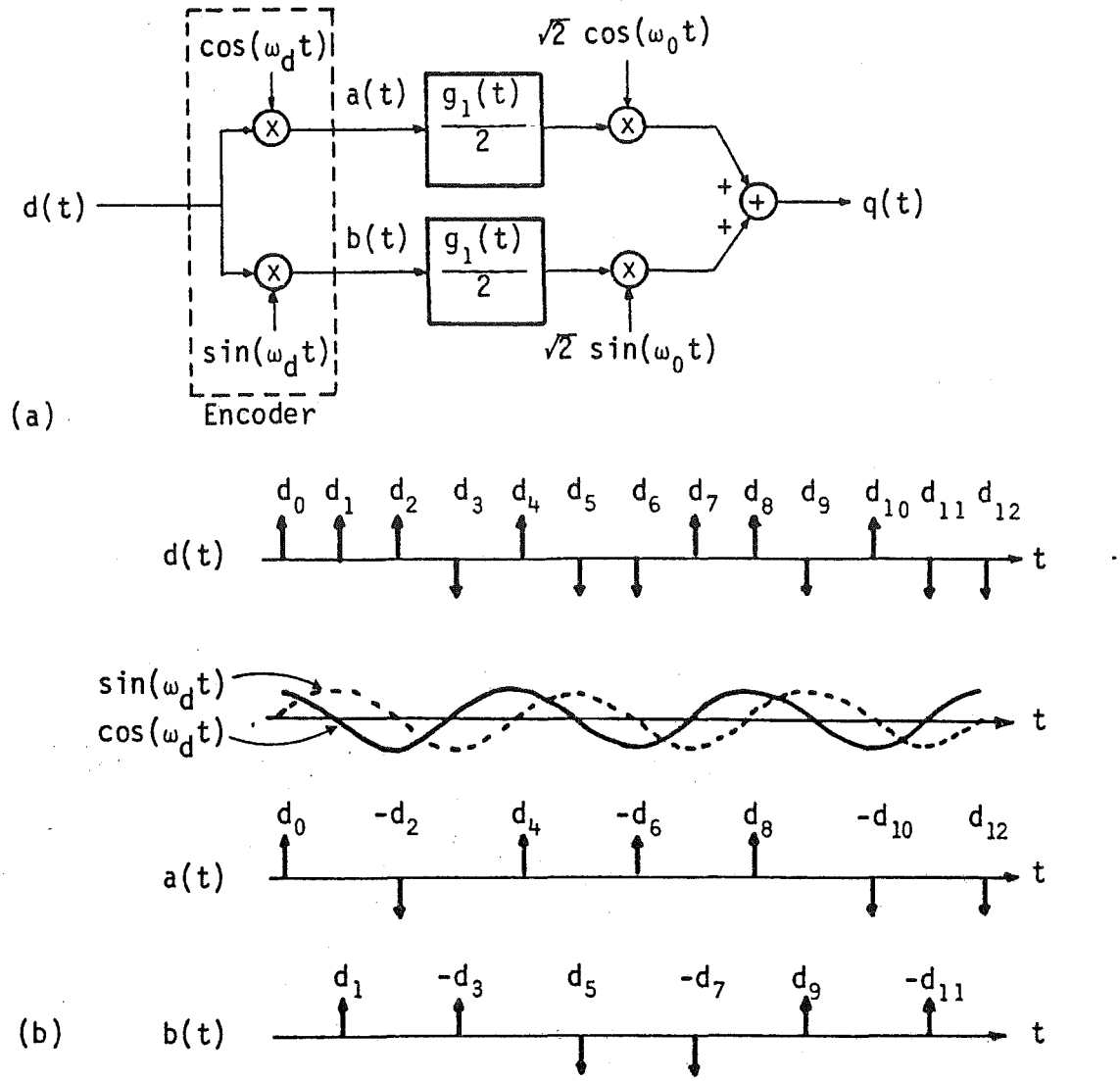


Fig. 3.13 (a) VSB generation using OQPSK techniques.
(b) Encoding of information bits.

$$x'(t) = [\{q'(t) * h_2(t)\}\ \sqrt{2}\cos(\omega_1 t)] * \ell(t)$$

$$= \left[\mathrm{Re}\left[q'(t) * \left\{g_2(t)e^{j\omega_0 t}\right\}\right] \times \mathrm{Re}\left[\sqrt{2}\, e^{j\omega_1 t}\right]\right] * \ell(t) \qquad (3.57)$$

Defining $x'(t)$ relative to ω_0 gives

$$x'(t) = \left[\mathrm{Re}\left[q'(t) * \left\{g_2(t)\, e^{j\omega_0 t}\right\}\right] \times \mathrm{Re}\left[\sqrt{2}\, e^{-j\omega_d t}\, e^{j\omega_0 t}\right]\right] * \ell(t)$$

$$= \frac{\sqrt{2}}{2}\,\mathrm{Re}\left[\left\{\left(q'(t)\, e^{-j\omega_0 t}\right) * g_2(t)\right\} e^{j\omega_d t}\right]$$

$$= \left[\{q'(t)\ \sqrt{2}\cos(\omega_0 t)\} * \frac{g_2(t)}{2}\right]\cos(\omega_d t)$$

$$+ \left[\{q'(t)\ \sqrt{2}\sin(\omega_0 t)\} * \frac{g_2(t)}{2}\right]\sin(\omega_d t) \qquad (3.58)$$

Equ. (3.58) describes the model shown in Fig. 3.14(a). If the zero crossings of $\cos(\omega_d t)$ and $\sin(\omega_d t)$ are aligned with the sampling times for $x'(t)$, then the sampling points can be relocated in receiver as shown in Fig. 3.14(b). This is the form of an OQPSK receiver.

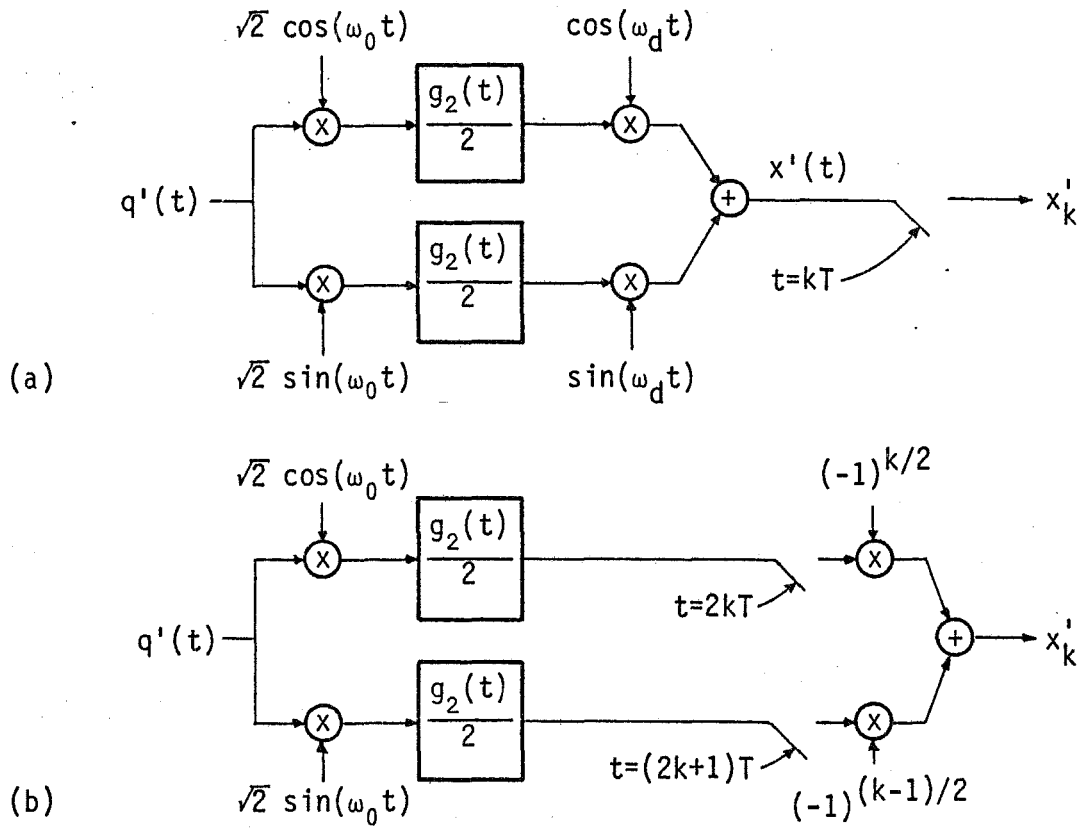


Fig. 3.14 (a) VSB demodulation using quadrature related paths. (b) VSB demodulation using OQPSK-type demodulation and sampling.

3.6 A 9600 BITS/SEC DUPLEX DATA SYSTEM

A 9600 bits/sec duplex data system intended for out-of-voiceband transmission over the subscriber telephone loop is described. The development of the system did not, however, reach hardward completion due to implementation difficulties which were encountered. It was these difficulties which initiated the more detailed study of digital transmission systems, in particular VSB.

The subscriber telephone loop enjoys high penetration into the private and business sectors, so it is understandable from an economic point of view that this medium is an attractive link for information transfer to or from these sectors. The subscriber loop was originally intended for analogue voice transmission, but it is now coming under increasing use as a medium for conveying digital information. If analogue voice and digital information are required to co-exist on the one subscriber loop, the digital information can be placed in a higher frequency band than the voice which requires the lower 4 kHz.

The basic design considerations for the 9600 bits/sec data system operating over the subscriber loop were:

(a) Duplex transmission
(b) Data and analogue voice transmission transparent to each other
(c) A BER not exceeding 10^{-5}
(d) Signal attenuation in the transmission link may be up to 50 dB.

There are several methods of achieving duplex transmission over the subscriber loop:

(a) Physical separation using a 4-wire circuit

(b) Time separation of the two signalling directions

(c) Cancellation of unwanted signals using a hybrid or echo canceler

(d) Frequency separation.

The cost of outside plant inhibits the use of a 4-wire link unless it already exists, so this alternative was not considered. Time separation using a method called burst mode transmission has been used to provide 64 kbits/sec duplex transmission of PCM voice [Bowman, et al., 1978; Inoue, 1979; Meyer and Roste, 1979; Torbergsen and Meyer, 1978], and even higher rates when data plus PCM voice are simultaneously transmitted [Andry, et al., 1981; Bylanski, et al., 1980]. Buckrell [1980], and Holte and Stuefflotten [1981] describe the use of an adaptive echo canceler to achieve duplex transmission at 80 kbits/sec. These latter two techniques require a relatively high degree of transmitter and receiver complexity and, as Dorros [1983] indicates, frequency multiplexing for data rates in the order of 8 kbits/sec can be used to provide duplex transmission.

With frequency multiplexing as the method chosen for the duplex transmission, it is preferable to keep all data signal passbands above 20 kHz to avoid any mutual interference with the analogue voice. From New Zealand Post Office survey information [NZPO, Oct. 1977] the average length of a subscriber loop in New Zealand is 2.35 km, with a standard deviation of 2.32 km, so the majority of subscribers will be served by cables less than 6 km in length. The cables vary both in make-up and size. Assuming the norm to be 0.5 mm PEUT, the NZPO transmission simulation programme, POTRANZ, indicates 50 dB attenuation at 6 km corresponds to an upper frequency limit of 120 kHz. To avoid marginal design the signal passbands should be kept well below this frequency.

The basic model for the duplex data plus voice system is shown in

Fig. 3.15. To avoid excessive crosstalk between the duplex data channels they must be separated by an adequate stop-band. Mutual interference can also be caused by system nonlinearities that may be present, so the two data passbands should not be harmonically related. A spectrally efficient system with good BER performance is also required. Following previous research by Lowe and Webb [1980], VSB using in-band generation techniques was proposed as a suitable efficient and relatively simple transmission method.

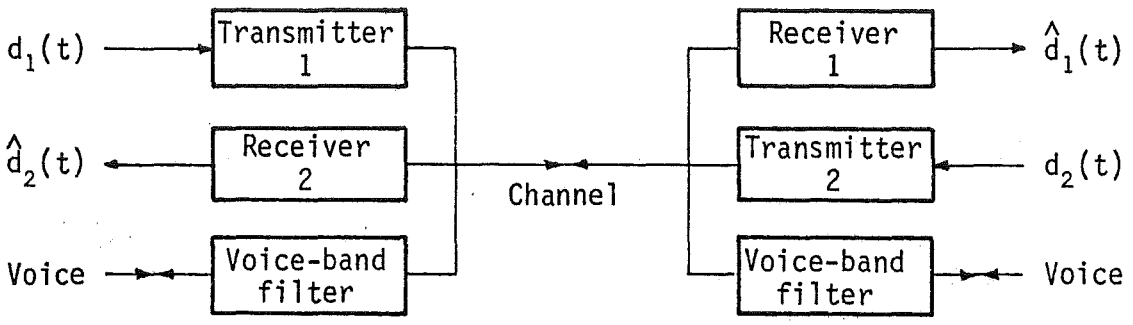


Fig. 3.15 Basic structure of voice plus duplex data system.

Analogue bandpass filters were used at the transmitter and receiver for VSB shaping and to avoid space consuming inductors in the circuitry, active filtering stages were employed. Carrier frequencies of

$$f_1 = 6 \times 9.6 \text{ kHz}$$

$$= 57.6 \text{ kHz} \qquad (3.59)$$

and

$$f_2 = 4 \times 9.6 \text{ kHz}$$

$$= 38.4 \text{ kHz} \qquad (3.60)$$

were chosen to give passband separation and to avoid excessive signal attenuation due to the cable characteristics.

The design and construction of the circuitry for the VSB path with carrier frequency 57.6 kHz was carried out first. The system stages were as shown in Fig. 3.1, except in-band generation was used at the transmitter. Signal shaping was divided equally between transmitter and receiver, with both filtering stages comprising of two pairs of identical butterworth bandpass filters. Stagger-tuning of the filter pairs enabled the required bandwidth to be met with steep filter skirts. Computer simulation was used to arrive at a suitable filtering configuration. The simulated and actual filter passbands are shown in Fig. 3.16. As a design aid, the criterion for zero ISI (see section 2.3.5) was used to evaluate and make a final decision on the filtering parameters that would be used. Defining the overall VSB passband frequency characteristic as H(f), then the tests for adherence to the zero ISI criterion were carried out on the baseband filter

$$H_c(f) = \frac{1}{2}[H(f - 57600) + H(f + 57600)]\,L(f). \qquad (3.61)$$

L(f) is included in (3.61) as a lowpass filter to indicate $H_c(f)$ does not contain any sum frequencies. The test results for the signal passband finally chosen are shown in Fig. 3.17. Following hardware construction, the eye-pattern in Fig. 3.18 was obtained, which illustrates how closely the zero ISI criterion was adhered to.

Some difficulties were encountered while developing the VSB system. Considerable effort, even using computer simulation, was required before the analogue filter parameters were finally decided on. Following this, the tolerance of the components used in constructing the filters meant there was no guarantee these filters would match simulation results.

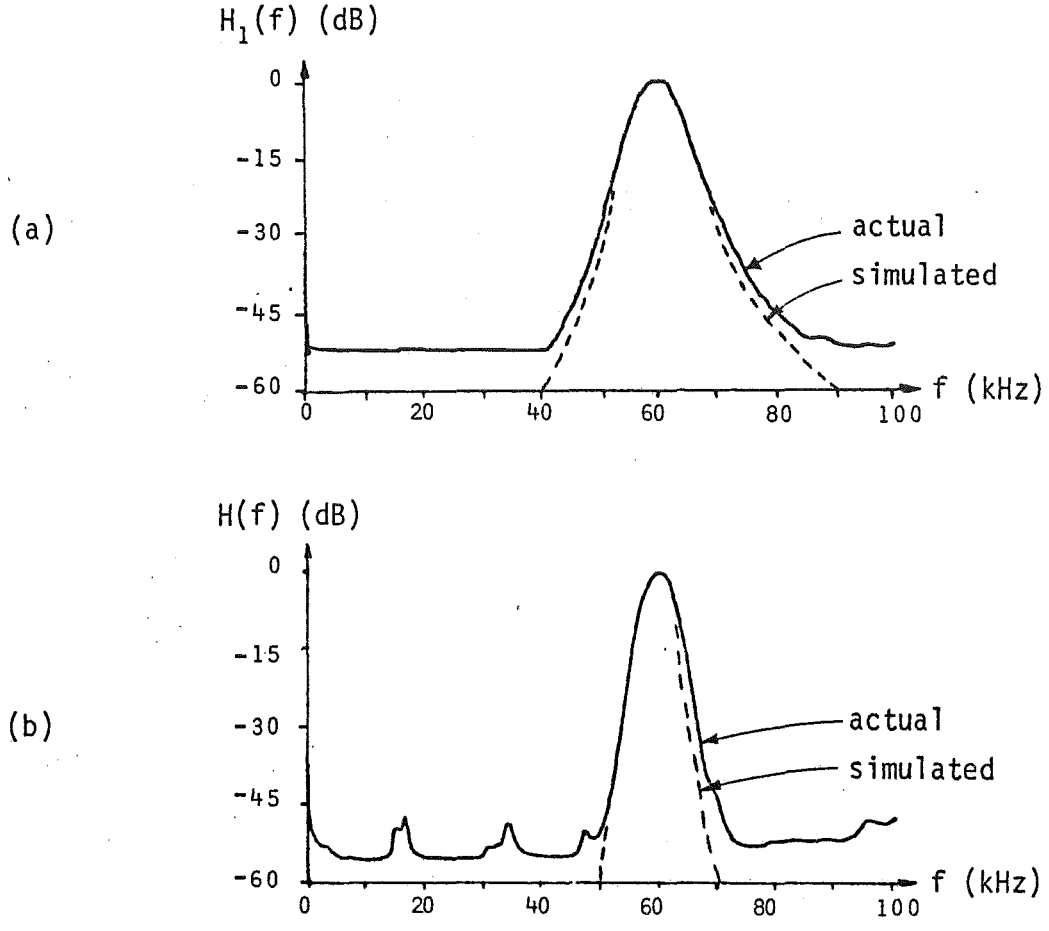


Fig. 3.16 (a) Transmitter and (b) overall passbands for 9600 bits/sec VSB system.

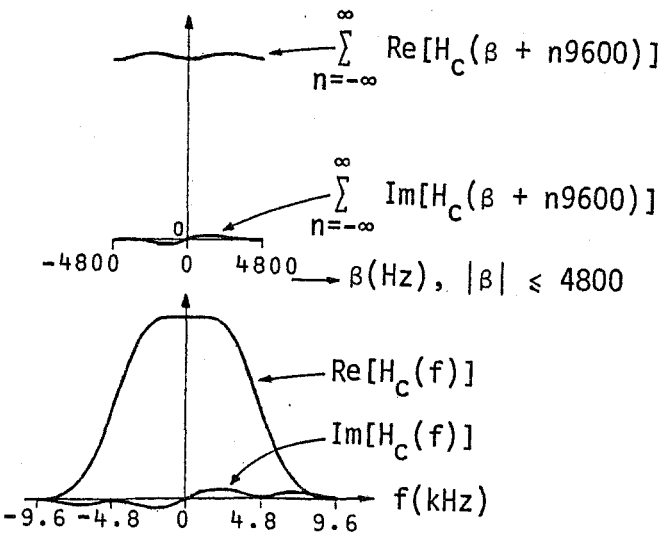


Fig. 3.17 Results of test for adherence of simulated passband to zero ISI criterion.

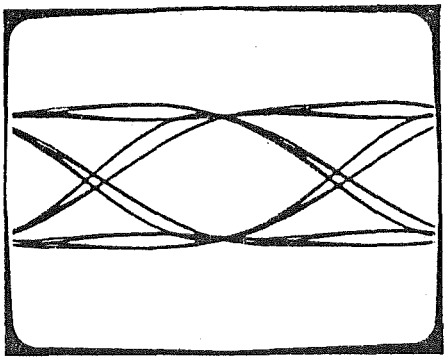
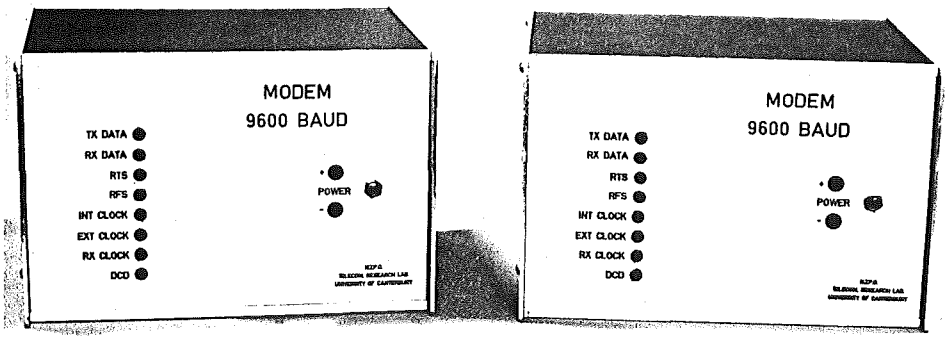


Fig. 3.18 Oscilloscope tracing of experimental VSB system eye-pattern.

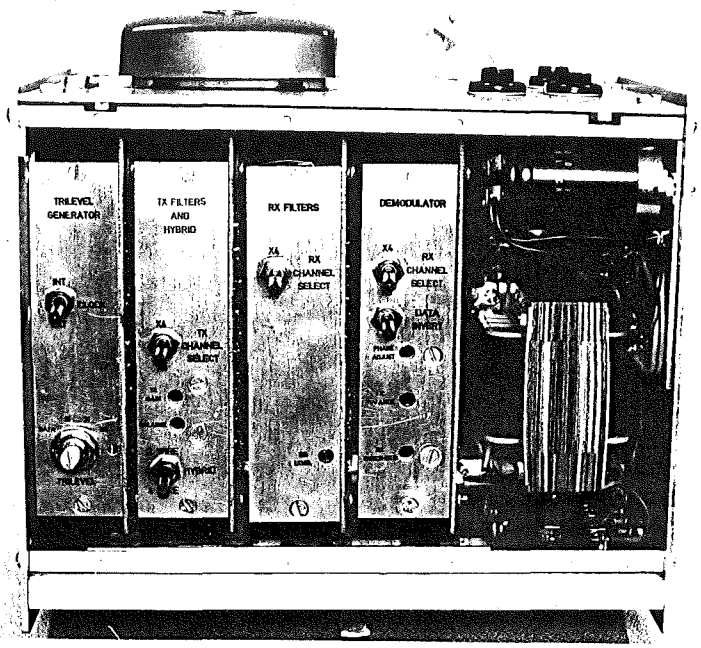
Adjustment of the resistors and capacitors in the filter circuitry was required following initial construction. So, from the initial simple concepts, actual design and implementation of the VSB system was complicated by the difficulty in defining the required signal passband using analogue filters. Also, the method of synchronization was unclear at the outset of the system development and the options available required further study. This latter point is the subject of the next chapter.

Photographs of the hardware constructed for the experimental VSB system are shown in Fig. 3.19.

(a)



(b)



(c)

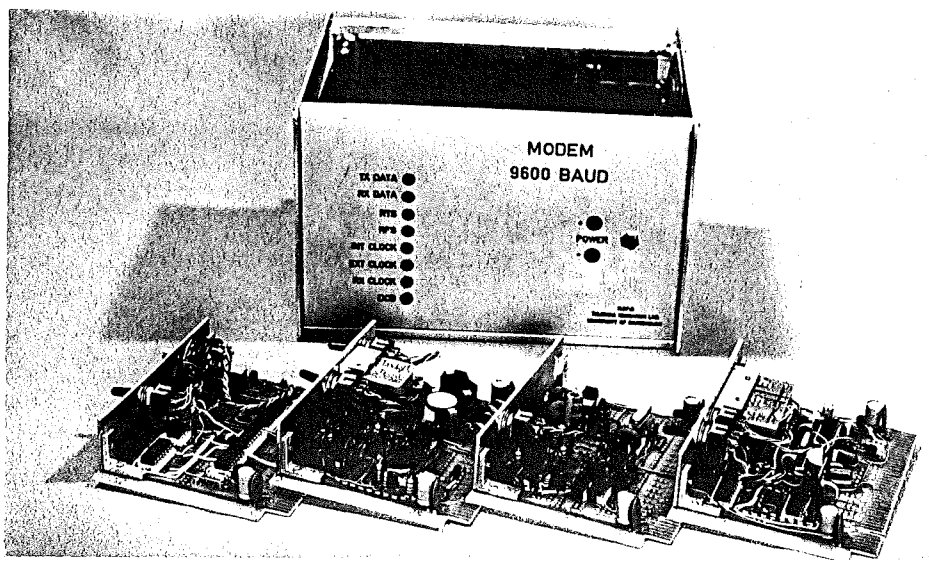


Fig. 3.19 Hardware constructed for the experimental VSB 9600 bits/sec duplex data system. (a) The two modems. (b) Top view of one modem. (c) The four circuit boards for one modem.

CHAPTER 4

SYNCHRONIZING VSB DATA TRANSMISSION

4.1 INTRODUCTION

In digital data transmission, synchronization is a crucial factor in determining system performance. Assuming a carrier-modulated system is involved, there are two levels of synchronization; namely, carrier and bit synchronization. First a reference carrier must be generated at the receiver for coherent demodulation of the received signal. A receiver clock must then be synchronized for sampling the baseband symbol sequence. Even when the incoming symbols are non-binary this latter operation is usually referred to as bit synchronization.

DSB offers the simplest method of carrier-modulated transmission, including carrier synchronization. Unfortunately, it is the least spectrally efficient since information is duplicated in each sideband. VSB is a compromise between DSB and SSB by allowing use of filters with gradual rolloff characteristics while retaining a reasonable level of spectral efficiency. However, use of VSB creates difficulties with carrier synchronization. In section 4.2 the reason for these difficulties is reviewed, along with schemes which have been proposed to overcome the problem.

Bit synchronization is a requirement independent of the carrier modulation technique used. However, as will be shown in section 4.3, the difficulty in deriving bit timing information from a baseband signal is very similar to that of deriving carrier information from VSB. Without the transmission of additional synchronization information, both carrier and bit timing waveforms must be derived from the received data signal. Data sequences can occur which result in the received VSB signal containing

no information for carrier or bit synchronization. Any data sequence restriction to avoid loss of one level of synchronization is seen to be in conflict with maintaining the other level of synchronization. In section 4.4 a new method of ensuring both levels of synchronization in a VSB system is described, where low-level pilot tones are added to the VSB signal to maintain synchronization under worst-case data sequence conditions.

4.2 REVIEW OF VSB CARRIER SYNCHRONIZATION

The general form of a VSB signal after receiver filtering was derived in Chapter 3:

$$v(t) = \left[\left\{d(t) * \frac{h_c(t)}{2}\right\} \sqrt{2}\cos(\omega_1 t)\right] + \left[\left\{d(t) * \frac{h_s(t)}{2}\right\} \sqrt{2}\sin(\omega_1 t)\right], \tag{4.1}$$

where $h_c(t)$ and $h_s(t)$ are baseband filter impulse responses, and d(t) is given by

$$d(t) = \sum_k d_k\, \delta(t - kT). \tag{4.2}$$

Binary transmission will be assumed, so $d_k = \pm 1$. The first term in (4.1) is the form of a DSB signal. Typically, $h_c(t)$ has a Fourier transform shown in Fig. 4.1(a). The second term modifies the DSB characteristics to that of a VSB signal. Fig. 4.1(b) shows an example of $H_s(f)$, the Fourier transform of $h_s(t)$. The resulting VSB amplitude spectrum is shown in Fig. 4.1(c).

Demodulation of v(t) by

$$c(t) = \sqrt{2}\cos(\omega_1 t + \theta_e(t)), \tag{4.3}$$

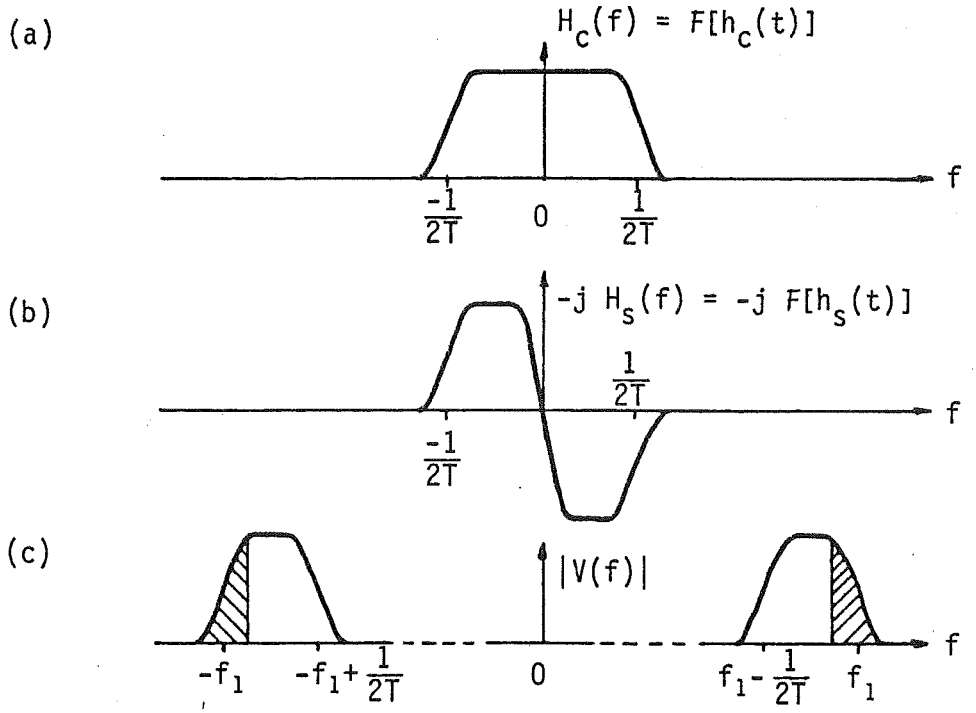


Fig. 4.1 Examples of (a) baseband filter spectrum for DSB shaping, (b) baseband filter spectrum for modifying DSB shaping to VSB, and (c) amplitude spectrum of VSB signal.

where $\theta_e(t)$ is the carrier phase error, results in a baseband signal

$$x(t) = \left[\left\{d(t) * \frac{h_c(t)}{2}\right\} \cos(\theta_e(t))\right] - \left[\left\{d(t) * \frac{h_s(t)}{2}\right\} \sin(\theta_e(t))\right]. \tag{4.4}$$

The second term in (4.4) is unwanted interference which can only be removed by having an accurate carrier reference; i.e., $\theta_e(t) = 0$.

Carrier recovery from the received VSB signal requires the presence of a fixed spectral component (there may be more than one) which has some harmonic relationship with the carrier frequency. The components can be isolated using narrow bandpass filtering [Franks, 1980] or phase locked loop (PLL) circuitry [Gardner, 1979]. The filtering or PLL stage effectively performs a time average over a long period of the signals within the processing bandwidth.

If an attempt is made to derive a reference carrier directly from $v(t)$ by passing it through a narrow bandpass filter $h_f(t)$, centered about f_1 Hz, the output of the filter will be a time averaged signal given by

$$v(t) * h_f(t) = K_{1f} E\left[d(t) * \frac{h_c(t)}{2}\right] \cos(\omega_1 t) + K_{1f} E\left[d(t) * \frac{h_s(t)}{2}\right] \sin(\omega_1 t), \tag{4.5}$$

where K_{1f} is a constant. If $d(t)$ is assumed to be a random sequence, then

$$E\left[d(t) * \frac{h_c(t)}{2}\right] = E\left[d(t) * \frac{h_s(t)}{2}\right]$$

$$= 0, \tag{4.6}$$

so

$$v(t) * h_f(t) = 0 \tag{4.7}$$

and no carrier can be recovered directly from $v(t)$.

One way to overcome this problem is to add a pilot tone at the carrier frequency to the VSB signal prior to transmission. However, since a data sequence may occur where $v(t)$ contains significant power at its carrier frequency for a prolonged period of time, the pilot tone must be of sufficient power (greater than the VSB signal) to ensure it is not cancelled or even reversed in phase. This method is therefore wasteful of available signalling power resulting in at least a 3 dB degradation in E_b/N_o performance. Also, note from (4.5) that if the second term in this equation is non-zero for prolonged periods the phase of any received carrier will be altered by the quadrature component.

A method to avoid pilot interference from the quadrature component was proposed by Ho [1974]. A pilot carrier is transmitted at a power level equal to that of the data signal. By requiring the VSB passband to be flat

about the carrier frequency results in no power in the quadrature component in this region and its interference with the pilot carrier phase is avoided, thus eliminating any "artificial" phase jitter. Ho states that by using information contained in the quadrature signal this method enables tracking of up to 60 Hz phase jitter due to channel impairments.

Bennett and Davey [1965] outline a method whereby a pilot tone is added in phase quadrature to the suppressed carrier. The locally generated carrier at the receiver is synchronized with the phase of the quadrature pilot. The pilot's phase is, however, modified by the in-phase portion of the VSB signal. The recovery method relies on the transmitted data being random so that when a narrow band filter or PLL is used the phase variation of the local carrier is small. This phase variation will then be approximately in linear relation to the average value of the recovered binary data. This average value can therefore be used to phase modulate the local carrier to remove most of the phase variation.

To avoid any interference between the VSB signal and the carrier pilot, Becker et al. [1962] (later described by Bennett and Davey [1965]) developed a VSB system where all VSB power is removed from a narrow band around the pilot. The missing VSB signal is reinserted at the receiver by using a regenerative feedback circuit. The method allows the use of a low level pilot, but the regenerative feedback is sensitive to any signal distortion and it tends to magnify any signal impairments.

More than one pilot can be added to the VSB signal. Wozencraft and Jacobs [1965] describe a method whereby low level pilot tones are inserted at the upper and lower edges of the passband. Their relative relationship to the carrier frequency enables recovery of a carrier which follows frequency and phase shifts of the VSB signal.

The VSB passband contains a DSB portion about its carrier

frequency as indicated by the shaded area in Fig. 4.1(c). A common method of deriving a carrier from a DSB signal is a frequency doubling nonlinearity such as a squarer [Haykin, 1978; Oberst and Schilling, 1971]. The same method can be used with VSB by making use of the signal power in its DSB portion. If v(t) is passed through a square-law nonlinearity, then the output, ignoring the baseband term, is

$$w(t) = \left[\left\{d(t) * \frac{h_c(t)}{2}\right\}^2 - \left\{d(t) * \frac{h_s(t)}{2}\right\}^2\right] \cos(2\omega_1 t)$$

$$+ \left[\left\{d(t) * \frac{h_c(t)}{2}\right\}\left\{d(t) * \frac{h_s(t)}{2}\right\}\right] \sin(2\omega_1 t). \qquad (4.8)$$

Time averaging w(t) by passing it through a narrow bandpass filter $h_{2f}(t)$, centered about $2f_1$ Hz, gives [Franks, 1980]

$$w'(t) = w(t) * h_{2f}(t)$$

$$= K_{2f} E\left[\left\{d(t) * \frac{h_c(t)}{2}\right\}^2 - \left\{d(t) * \frac{h_s(t)}{2}\right\}^2\right] \cos(2\omega_1 t), \qquad (4.9)$$

where K_{2f} is a constant.

Equation (4.9) indicates that the output of the narrow band filter is a sinusoid at twice the carrier frequency with the desired phase, but with an amplitude proportional to the difference in power between the quadrature related signals in v(t). Signal w'(t) will have maximum amplitude when d(t) consists of unipolar data and the quadrature signal contains no power. If w(t) goes to zero then synchronization will fail. This condition exists when v(t) has no signal power in its DSB portion for a prolonged period; i.e., has the spectral characteristics of SSB where the quadrature related signals in v(t) have equal power. The worst-case condition for synchronization will therefore occur when d(t) consists of alternating data bits. So, without the addition of pilot tones, the use of the frequency doubling nonlinearity requires restrictions on allowable

data sequences to avoid synchronization failure.

The operation of frequency doubling can alternatively be considered in the frequency domain. Squaring v(t) is equivalent to convolving V(f) (Fig. 4.1(c)) with itself. Spectral components at $2f_1$ Hz will be present at the squarer output only if V(f) contains signal frequency components (including negative frequencies) which differ by $2f_1$ Hz. For VSB this requires signal components to be present in its DSB portion. As the vestige is reduced a decreasing portion of the VSB passband contributes to the $2f_1$ Hz component at the squarer output.

4.3 BIT SYNCHRONIZATION : THE CONFLICT WITH VSB CARRIER RECOVERY

A typical eye-pattern for a baseband signal is shown in Fig. 4.2. With optimum narrow band passband filtering the power in adjacent bits overlaps, but there remains a point of zero ISI during each bit period. Sampling at these points avoids degradation from ISI, but requires synchronization of the sampling times with the received signal.

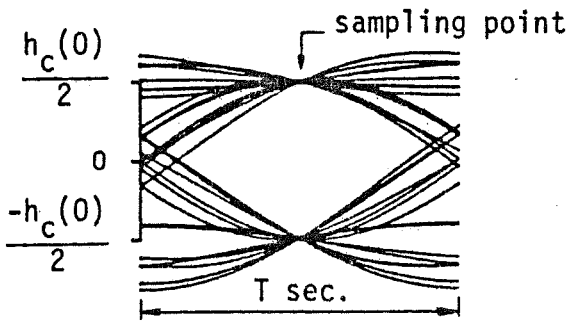


Fig. 4.2 Eye-pattern showing point of zero ISI.

An example of the amplitude spectrum of a baseband signal is shown in Fig. 4.3. For random data no fixed spectral component appears in X(f), so a bit timing signal cannot be derived directly [Bylanski and

Ingram, 1976] and nonlinear recovery techniques must be used [Bhargava et al., 1981]. These techniques rely on x(t) containing bit transitions, and additional timing bits may have to be inserted or code restrictions enforced to ensure that bit transitions occur [Bellamy, 1982].

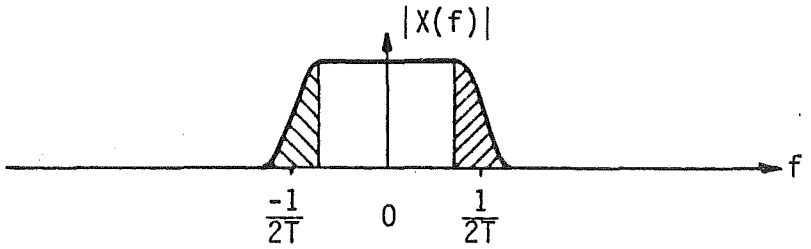


Fig. 4.3 Example of the amplitude spectrum of a baseband signal.

In a similar manner to VSB carrier recovery, a square-law nonlinearity can be used to derive bit timing information from x(t). If X(f) contains spectral components which differ by 1/T Hz then $x^2(t)$ will contain a sinusoidal component at 1/T Hz. The recovery method relies on signal power being present in the shaded area of X(f) shown in Fig. 4.3. If this shaded area is considered as a DSB portion about 1/(2T) Hz, this operation is identical to VSB carrier recovery using a square-law nonlinearity, except for the different frequencies involved.

A conflict is seen to exist between data sequences best suited for carrier and bit timing recovery for a VSB system when no pilot tones are transmitted. Reliable carrier synchronization is obtained by ensuring sufficient signal power is present within the DSB portion of the VSB signal. The best-case condition occurs when unipolar data are transmitted. Carrier synchronization will fail when alternating data are transmitted for any extended period. The converse is true for bit synchronization, where the worst-case situation occurs when d(t) consists of unipolar data, resulting in the signal power in x(t) being concentrated at very low frequencies. Code restrictions to avoid failure in one of the two levels

of synchronization must be to the detriment of the other level. It is therefore preferable to avoid any code restrictions by using carrier and bit synchronization schemes which are independent of the data sequence. For VSB this requires the use of pilot tones and to avoid significant E_b/N_o degradation any pilot tone should be at a low level. Such a scheme is presented in the following section.

4.4 PROPOSED VSB SYNCHRONIZATION SCHEME

A new VSB synchronization method is presented which enables both carrier and bit synchronization to be maintained independent of the data sequence [Henderson and Webb, 1984]. The method involves adding two pilot tones to the VSB signal; one at the carrier frequency to maintain carrier synchronization and the other 1/(2T) Hz from the carrier so that on demodulation the baseband signal contains a pilot at 1/(2T) Hz which can maintain bit synchronization. These pilot tones are kept at a low level to minimise the E_b/N_o degradation due to them consuming available channel signal power and also distorting the recovered baseband signal.

4.4.1 General Model

The model of the VSB system with synchronization is shown in Fig. 4.4. The cascade of $h_1(t)$ and $h_2(t)$ determine the VSB passband shaping, where

$$h(t) = h_1(t) * h_2(t) \tag{4.10}$$

is the overall shaping function.

AWGN n(t) with two-sided spectral density $N_o/2$ is assumed to be the only channel impairment.

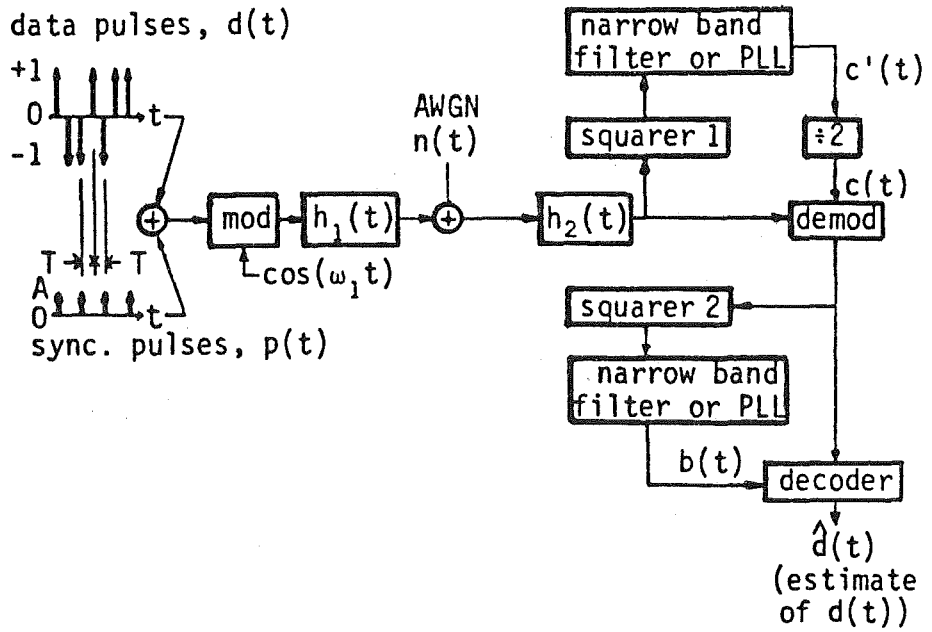


Fig. 4.4 Model for carrier and bit synchronization of VSB.

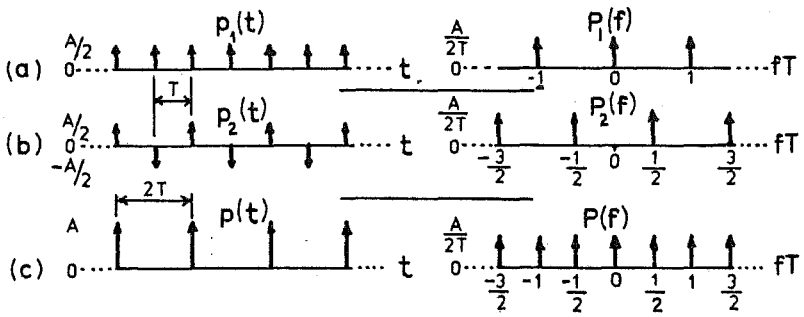


Fig. 4.5 Time and frequency domain characteristics of synchronization pulse sequences for (a) carrier synchronization, (b) bit synchronization, and (c) a combination of (a) and (b).

Pilot tones are added to the VSB signal by way of synchronization pulses summed with d(t). With reference to Fig. 4.5, pulse sequence $p_1(t)$ when summed with d(t) would insert a carrier pilot in the VSB signal, while the addition of $p_2(t)$ would insert a pilot 1/(2T) Hz from the carrier. The summation of $p_1(t)$ and $p_2(t)$ gives p(t), the actual sequence added to d(t). The resulting VSB and baseband amplitude spectra will now include low level pilots as shown in Fig. 4.6.

The pilot at f_1 ensures presence of power for carrier synchronization if d(t) is such that v(t) takes on the spectral characteristics of an SSB signal. The pilot at $(f_1 - 1/(2T))$ is translated down to baseband (Fig. 4.6(b)) and ensures bit timing information is present when x(t) has its power concentrated about d.c.

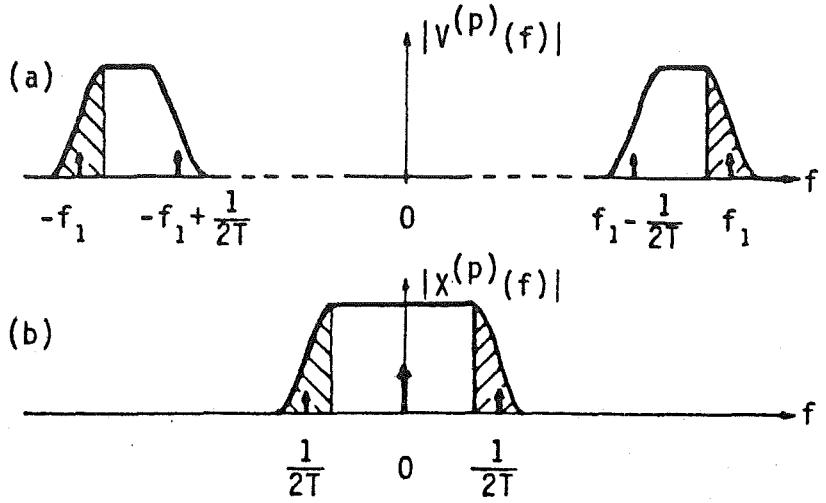


Fig. 4.6 Amplitude spectra of (a) VSB and (b) baseband signals with synchronization pilots.

The carrier and bit timing recovery stages use squarers followed by narrow bandpass filters or PLL's to isolate the desired spectral component; either

$$b(t) = \cos\left[\frac{2\pi}{T}(t + \tau)\right] \tag{4.11}$$

for bit timing, where τ is the sampling error, or

$$c'(t) = \cos[4\pi f_1 t + 2\theta_e(t)] \tag{4.12}$$

for carrier synchronization, where $\theta_e(t)$ is the phase error in the recovered carrier c(t) after frequency dividing by 2.

4.4.2 BER Performance

The BER performance is derived to determine the combined degradation due to the pilots consuming available channel power and their distorting effect on the recovered baseband signal.

Perfect carrier synchronization is assumed which simplifies analysis by enabling use of the equivalent baseband model shown in Fig. 4.7. The transmitter and receiver filters have respective impulse responses $g_{c1}(t)$ and $g_{c2}(t)$, where

$$\frac{h_c(t)}{2} = g_{c1}(t) * g_{c2}(t) \tag{4.13}$$

is the overall symbol shaping function. With matched filtering

$$g_{c1}(t) = g_{c2}(-t) \tag{4.14}$$

and for zero ISI

$$h_c(kT) = h_c(0)\ \delta_{ko}\ , \quad k \text{ an integer.} \tag{4.15}$$

The input to the decoder will be

$$x'(t) = x^{(p)}(t) + \alpha(t), \tag{4.16}$$

where

$$x^{(p)}(t) = [d(t) + p(t)] * \frac{h_c(t)}{2}$$

$$= d^{(p)}(t) * \frac{h_c(t)}{2} \tag{4.17}$$

and

$$\alpha(t) = n(t) * g_{c2}(t). \tag{4.18}$$

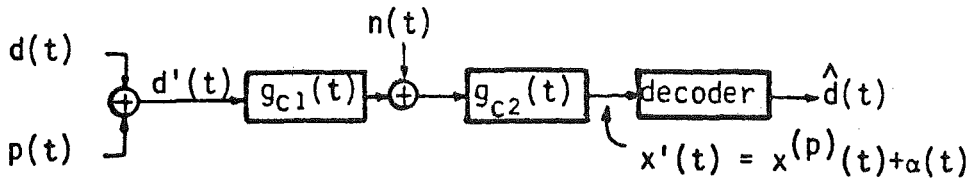


Fig. 4.7 Equivalent baseband model when perfect carrier synchronization is assumed.

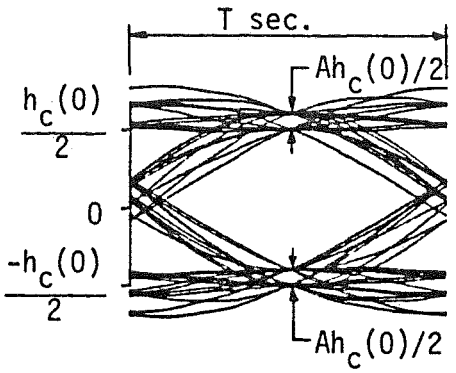


Fig. 4.8 Eye-pattern showing distortion due to the addition of synchronization pilots.

The system eye-pattern is shown in Fig. 4.8. Since zero ISI is assumed, p(t) will modify the sampled signal for every second received bit. The decoder used is the same as for baseband antipodal signalling. It simply makes a zero threshold decision at each sampling point on whether a positive or negative level was received. In section 2.4 it was shown that for a sampled level

$$x_k' = x_k^{(p)} + \alpha_k, \tag{4.19}$$

the probability that noise α_k causes x_k' to have opposite polarity to $x_k^{(p)}$ (i.e., a decoding error occurring) is

$$P_e = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{|x_k^{(p)}|}{\sqrt{2}\,\sigma_\alpha}\right), \tag{4.20}$$

where

$$\sigma_\alpha^2 = \frac{N_o}{2} \int_{-\infty}^{\infty} |G_{c2}(f)|^2 \, df$$

$$= \frac{N_o \, h_c(0)}{4} \tag{4.21}$$

is the power of the filtered AWGN. Using (4.20) and assuming random data, the BER for the system shown in Fig. 4.7 will be

$$P_b = \frac{\operatorname{erfc}\left[\frac{h_c(0)/2}{\sqrt{2}\,\sigma_\alpha}\right] + \tfrac{1}{2}\operatorname{erfc}\left[\frac{(1+A)\,h_c(0)/2}{\sqrt{2}\,\sigma_\alpha}\right] + \tfrac{1}{2}\operatorname{erfc}\left[\frac{(1-A)\,h_c(0)/2}{\sqrt{2}\,\sigma_\alpha}\right]}{4} . \tag{4.22}$$

The energy per bit in the channel is given by

$$E_b = E[d^{(p)}(t)] \int_{-\infty}^{\infty} |G_{c1}(f)|^2 \, df$$

$$= \frac{(1+A)^2 + (1-A)^2 + 2}{4} \, \frac{h(0)}{2}$$

$$= (1 + A^2/2) \, \frac{h(0)}{2} . \tag{4.23}$$

Substituting (4.21) and (4.23) into (4.22) gives

$$P_b = \frac{1}{4}\operatorname{erfc}\left[\left\{\frac{E_b}{(1+A^2/2)N_o}\right\}^{\frac{1}{2}}\right] + \frac{1}{8}\operatorname{erfc}\left[(1+A)\left\{\frac{E_b}{(1+A^2/2)N_o}\right\}^{\frac{1}{2}}\right]$$

$$+ \frac{1}{8}\operatorname{erfc}\left[(1-A)\left\{\frac{E_b}{(1+A^2/2)N_o}\right\}^{\frac{1}{2}}\right] \tag{4.24}$$

Plots of P_b for different values of A are shown in Fig. 4.9.

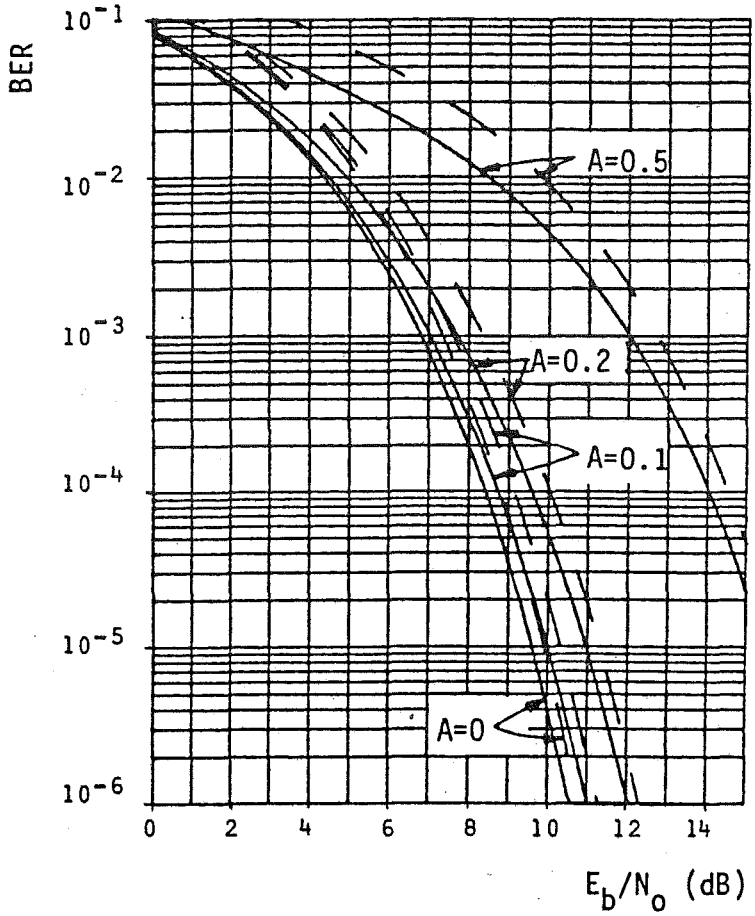


Fig. 4.9 BER curves showing degradation due to synchronization pilots of amplitude A. Dashed lines are for differential encoding.

The proposed carrier recovery method results in a 180 degree phase ambiguity due to the frequency divide-by-2 stage. The ambiguity can be overcome using differential encoding. This results in a degradation in BER performance relative to P_b according to the relationship [Bhargava, et al., 1981]

$$P_{de} = 2P_b(1 - P_b). \tag{4.25}$$

Plots of P_{de} are shown by the dashed lines in Fig. 4.9.

4.4.3 Simulation Model and Results

Computer simulation was carried out to obtain estimates of expected phase errors in carrier and bit timing recovery. Of main interest are the worst-case situations where the pilot tones must maintain carrier and/or bit synchronization without the aid of the data signal.

Simulation can be simplified by using the similarity between carrier and bit timing recovery that was noted in section 4.3. This enables the equivalent baseband model shown in Fig. 4.10 to be used. Results for bit timing are obtained directly. These results can be equated to carrier recovery since the only difference is the frequency of the recovered waveform. The worst-case bit timing recovery situation, which occurs when d(t) consists of unipolar data bits, can be equated to the worst-case carrier recovery situation, which occurs when d(t) consists of alternating bits.

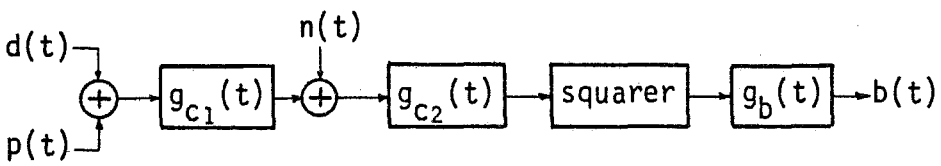


Fig. 4.10 Simulation model used to obtain plots of b(t).

The transmitter and receiver filters in Fig. 4.10 conform with equations (4.13) - (4.15). Results were obtained for $h_c(t)$ having lowpass raised-cosine filtering characteristics. Its -6 dB bandwidth was fixed at 1/(2T) Hz while rolloff factors 0.5, 0.25 and 0.15 were tried. Note that if ρ_{vsb} is the rolloff factor of the VSB shaping filter and ρ is the rolloff factor of the shaping filter for the equivalent baseband system, then

$$\rho_{vsb} = 2\rho. \tag{4.26}$$

For $g_b(t)$, raised-cosine filters with unity rolloff factor and -6 dB bandwidths of either B = 1/(200T) or B = 1/(500T) were used. The level of the synchronization pulses was set at A = 0.2. In terms of synchronization at the receiver A should be as large as possible; however, it is also desirable to avoid excessive E_b/N_o degradation. The chosen value of A is a compromise between these two requirements. A BER performance of 10^{-5} is acceptable for most applications, and with differential encoding this corresponds to an E_b/N_o value of approximately 11.5 dB (see Fig. 4.9). E_b/N_o was set at this level for all simulation runs.

A convenient method of viewing variations in the recovered bit timing waveform b(t) is the "eye-pattern" representation [Franks and Bubrouski, 1974]. Successive cycles of b(t) are plotted on the same time base which has a duration equal to the bit period. The results for several runs with different parameters are shown in Fig. 4.11. Although not illustrated, for the case where A = 0 and d(t) consists of random data the displays for b(t) showed no noticeable difference from those for random data and A = 0.2 in Fig. 4.11.

It is the zero crossings of b(t) which contain the synchronization information. The "eye-patterns" in Fig. 4.11 have been normalised to the same peak amplitude for the purpose of display. The zero crossing information is not affected by this, and in practice b(t) would usually be passed through a hardlimiter stage to remove amplitude fluctuations. The zero crossings in b(t) would be used to synchronize sampling times, so from the displays a measurement of the peak zero crossing variation from the mean corresponds to the peak sampling error in bit synchronization. Results from several simulation runs, including those illustrated in Fig. 4.11, are given in Table 4.1. Also given in this table are estimates of the peak phase error for the recovered carrier component c(t). The tabulated phase errors were obtained by first equating zero crossing

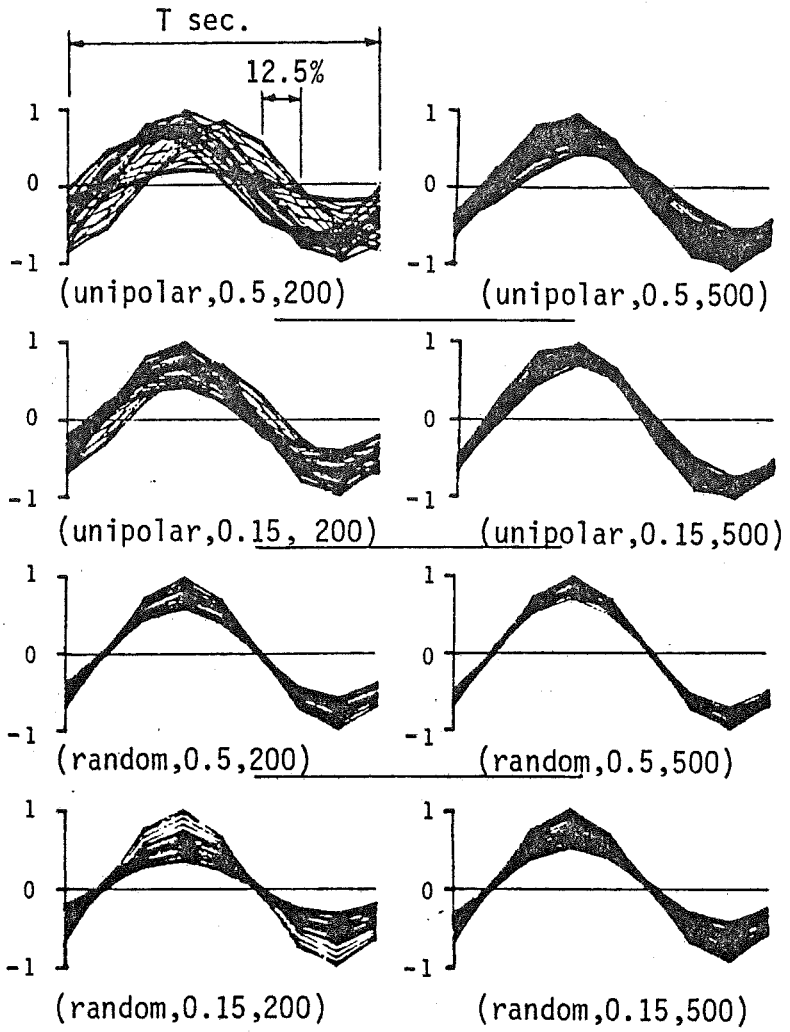


Fig. 4.11 "Eye-patterns" of simulated bit timing waveforms for indicated parameters (unipolar or random data, ρ, 1/(BT)). E_b/N_o = 11.5 dB and A = 0.2. Amplitude of plots normalised.

variations in b(t) to phase variations in the twice-carrier component c'(t). The frequency divide-by-2 stage will halve this phase variation to that indicated for c(t). As an example, for d(t) a unipolar data sequence, ρ = 0.5 and BT = 1/200, Table 4.1 shows the measured peak variation in zero crossings as ±12.5%. This can be equated to a ±45 degree phase variation in the twice-carrier component for identical parameters, except d(t) consisting of alternating data. This phase variation is then equal to ±22.5 degrees for the recovered carrier following frequency divide-by-2.

TABLE 4.1

SUMMARY OF SIMULATION RESULTS FOR VSB SYNCHRONIZATION

Data Sequence d(t)	ρ	1/(BT)	Measured Peak Zero Crossing Variation in b(t) (percent)	Estimated Peak Phase Error in c(t) (deg)
Unipolar	0.50	200	±12.50	0
"	0.50	500	± 6.25	0
"	0.25	200	±10.00	0
"	0.25	500	± 3.75	0
"	0.15	200	± 7.50	0
"	0.15	500	± 2.50	0
Alternating	0.50	200	0	±22.50
"	0.50	500	0	±11.25
"	0.25	200	0	±18.00
"	0.25	500	0	± 6.75
"	0.15	200	0	±13.50
"	0.15	500	0	± 4.50
Random	0.50	200	± 1.25	± 2.25
"	0.50	500	± 0.95	± 1.71
"	0.15	200	± 1.90	± 3.42
"	0.15	500	± 1.60	± 2.88

4.4.4 Discussion of Results

For bit synchronization, when unipolar data are transmitted no level transition occurs in the baseband signal and the pilot in the VSB signal at $(f_1 - 1/(2T))$ Hz must maintain bit synchronization by itself. Reference to Table 4.1 indicates that for 1/(BT) fixed, the zero crossing variations (i.e. noise interference) decreases with ρ. This is because noise at the output of $g_b(t)$ in Fig. 4.10 is due to noise at the input of

the squarer which is within the rolloff portion of the receiver filter's passband. Thus as ρ decreases, so does the effective noise bandwidth of the interference at the input to the squarer. A further reduction in zero crossing variation is achieved by reducing the bandwidth of $g_b(t)$ or, equivalently, the bandwidth of a PLL if it were to replace $g_b(t)$. Consideration must however be given to the time required to initialize the bit timing circuitry. For a PLL with bandwidth B Hz the time for settling is the same order of magnitude a 1/B seconds [Rich, 1974]. The above points also apply equally to carrier recovery, except the worst-case situation occurs for alternating data.

An indication of the effect of carrier phase error and different values of ρ_{vsb} is shown by the eye-pattern plots in Fig. 4.12. As ρ_{vsb} decreases, the width of the eye opening also decreases, hence so does the allowable error in the sampling times by the decoder [Huang et al., 1979]. Any carrier phase error emphasises the reduction in the eye openings for low values of ρ_{vsb}. The VSB system therefore becomes less tolerant of carrier phase and bit timing errors as ρ_{vsb} is reduced. Results show that this requirement is adhered to for both levels of synchronization where, for the worst-case situations shown in Table 4.1, errors become smaller as ρ decreases.

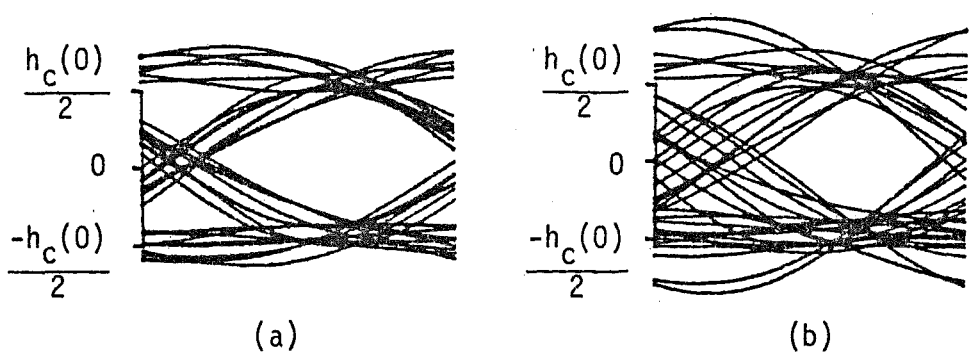


Fig. 4.12 VSB system eye-patterns when carrier phase error is 12 degrees and passband rolloff factors are (a) $\rho_{vsb} = 1$ and (b) $\rho_{vsb} = 0.3$.

For random data, the carrier phase and bit timing errors are seen to increase as ρ decreases. This is in conflict with the reduction in eye openings as ρ is reduced, but since these errors are so small this conflict is of little consequence. The two low level pilot tones have a negligible role in maintaining synchronization when the data are random, since for this situation the bulk of the power in the recovered synchronization waveforms is obtained from the data signal. The variation in phase errors with ρ agrees with (4.9), which indicates that the power in the twice-carrier component decreases as the vestigial sideband portion of the VSB passband is reduced. This also applies to bit timing, where the power of the recovered waveform decreases with ρ.

The presence of low level pilot tones is only necessary when worst-case data sequences occur. An adaptive system could be used, where synchronization pulses are transmitted only when a situation occurs where the data signal does not contain the information for synchronization. Since the worst-case situations for carrier and bit synchronization are generally mutually exclusive, an adaptive system could use pulse sequences $p_1(t)$ and $p_2(t)$ (Fig. 4.5) independently.

The addition of synchronization pulses with the data sequence distorts every second sample of the recovered baseband signal. This distortion is simply an amplitude shift by an amount proportional to the level of the synchronization pulses. If the "intelligence" of the receiver was increased, enabling the decoder to follow these amplitude shifts, the distorting effect of the synchronization pulses could be negated. The only E_b/N_o degradation due to the presence of the pilot tones would then be from them consuming available signal power in the channel. Using (4.20), (4.21) and (4.23), the BER performance would be given by

$$P_b = \tfrac{1}{2}\,\mathrm{erfc}\left[\left\{\frac{E_b}{(1+A^2/2)N_o}\right\}^{\frac{1}{2}}\right], \qquad (4.27)$$

so the E_b/N_o degradation due to the pilot tones would be 10 $\log_{10}(1+A^2/2)$ dB. This more complex decoding method would enable larger pilot tones to be used without suffering the excessive E_b/N_o degradation as for the simple zero-threshold decoder. However, the added complication to the receiver may not be justified, since the presence of synchronization pilots is really only required under worst-case data sequence conditions.

CHAPTER 5

TRANSMISSION OVER NONLINEAR CHANNELS

5.1 INTRODUCTION

The models and analysis of communication systems in the previous chapters have assumed the linear transmission of data signals between transmitter and receiver, with the only channel impairment as AWGN. In many situations repeaters must be used in the channels to maintain signals at an acceptable level above extraneous interference. Linear signal amplification at a repeater is preferable in terms of minimising signal distortion due to this stage; however, efficient utilisation of available power may require the repeater to be operated in a nonlinear amplification mode.

The satellite channel is an excellent example of where the power available at a repeater is limited. Most of the research into nonlinear transmission channels has been directed at satellite communications. The modulation formats which have received most attention are QPSK, OQPSK and MSK (minimum-shift keying) with constant signal envelopes that are suited to nonlinear channels. Current signalling techniques available for transmission over nonlinear channels are reviewed in section 5.2.

As with linear systems, spectral efficiency is still an important consideration in using a particular signalling format over a nonlinear channel. Often the approach to finding a suitable signalling format has been to start with a constant envelope signal and then attempt to find the optimum filter for bandlimiting the signal in a nonlinear channel. The alternative approach of starting with a spectrally efficient signal, namely

VSB, and determining the effects of hardlimiting is studied in section 5.3. Results indicate that VSB is a suitable candidate for nonlinear transmission.

5.2 REVIEW OF NONLINEAR TRANSMISSION

In this section the techniques commonly considered for transmission over nonlinear channels are reviewed; namely, QPSK, OQPSK and MSK. Also covered are two similar systems aimed at increased spectral efficiency; these are quadrature overlapped raised-cosine (QORC) [Austin and Chang, 1981] and intersymbol-interference and jitter-free offset QPSK (IJF-OQPSK, or Feher's QPSK) [Feher, 1981, 1983; Le Ngoc and Feher, 1983; Le Ngoc, et al., 1982]. These systems are compared with VSB in section 5.4.

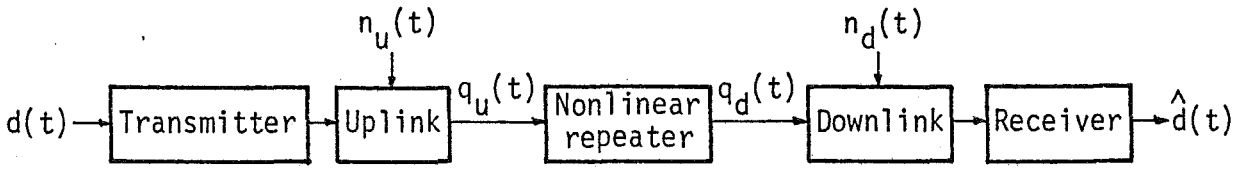


Fig. 5.1 General model for a nonlinear system.

A general model for a nonlinear system is shown in Fig. 5.1. Since most research in this area is concerned with satellite channels, the terms uplink and downlink will be used, respectively, to refer to the channel paths before and after the repeater. $n_u(t)$ and $n_d(t)$ are signal distorting terms usually assumed to consist of AWGN and possibly combinations of cochannel interference, adjacent channel interference and fading [Bhargava, et al., 1981; Fang, 1981; Feher, 1981; Kennedy and Shimbo, 1981].

The satellite transponder nonlinearity is usually a travelling wave tube amplifier (TWTA) [Strauss, et al., 1981]. A high power amplifier (HPA) may be used at the transmitter. Fang [1981] gives typical transfer

characteristics for both types of amplifiers. For maximum power efficiency the amplifiers will be operated near their saturation regions where they exhibit severe nonlinear characteristics. If a signal with the general form

$$q_u(t) = A(t) \cos(\omega_1 t + \theta(t)) \tag{5.1}$$

is input to a TWTA (or HPA), the output will be [Bhargava, et al., 1981]

$$q_d(t) = U(A) \cos(\omega_1 t + \theta(t) + \phi(A)), \tag{5.2}$$

where $U(A)$ is known as the AM-AM distortion and $\phi(A)$ as the AM-PM distortion. Note that if $A(t)$ is constant no distortion is introduced and $\phi(A)$ will simply be a fixed phase shift.

In simulation studies of nonlinear channels the nonlinearity is often considered to be a bandlimited hardlimiter which has readily defined characteristics. Schwartz, et al. [1966] consider the hardlimiter as a suitable benchmark against which other repeater configurations can be measured. To avoid the variation in the performance of a TWTA with input signal amplitude, Schwartz, et al. [1966] consider placing a hardlimiter before the TWTA to ensure the amplifier operates at a well chosen point.

The hardlimiter is always considered in conjunction with a bandpass function which passes only the fundamental components from the hardlimiter output. This follows the practical situation where signal harmonics at the output of nonlinear stages fall outside the bandwidth of the repeater circuitry. With $q_u(t)$ as the signal input to a bandpass hardlimiter, the output will have the form [Davenport and Root, 1958]

$$q_d(t) = K \cos(\omega_1 t + \theta(t)), \tag{5.3}$$

where K is a real constant. The bandpass hardlimiter produces no phase distortion, but if A(t) is not constant (i.e., the input signal has envelope variations), then AM-AM distortion will occur.

On the basis of QPSK, OQPSK (with rectangular pulse shaping) and MSK having constant envelopes, these three signalling formats have received considerable attention in the study of methods suitable for transmission over nonlinear channels. A general model for generating any of these signals is shown in Fig. 5.2. The differences between QPSK, OQPSK and MSK are determined by the symbol shaping used and the relative time positioning of the data sequences in the quadrature related paths. This is illustrated in Fig. 5.3. QPSK and OQPSK use identical symbol shaping given by

$$g_{1c}(t) = g_{1s}(t) = \begin{cases} 1, & |t| \leq T, \\ 0, & \text{elsewhere} \end{cases} \tag{5.4}$$

For OQPSK, the signal sequences $q_c(t)$ and $q_s(t)$ are offset relative to each other by T seconds. A similar offset applies to MSK but the symbol shaping is now sinusoidal and is given by

$$g_{1c}(t) = g_{1s}(t) = \begin{cases} \cos\left(\frac{\pi t}{2T}\right), & |t| \leq T, \\ 0, & \text{elsewhere.} \end{cases} \tag{5.5}$$

After modulation and summation, the signals will have the general form

$$\begin{aligned} q_u(t) &= q_c(t)\ \sqrt{2}\cos(\omega_1 t) + q_s(t)\ \sqrt{2}\sin(\omega_1 t) \\ &= A(t)\cos(\omega_1 t + \theta(t)), \end{aligned} \tag{5.6}$$

where

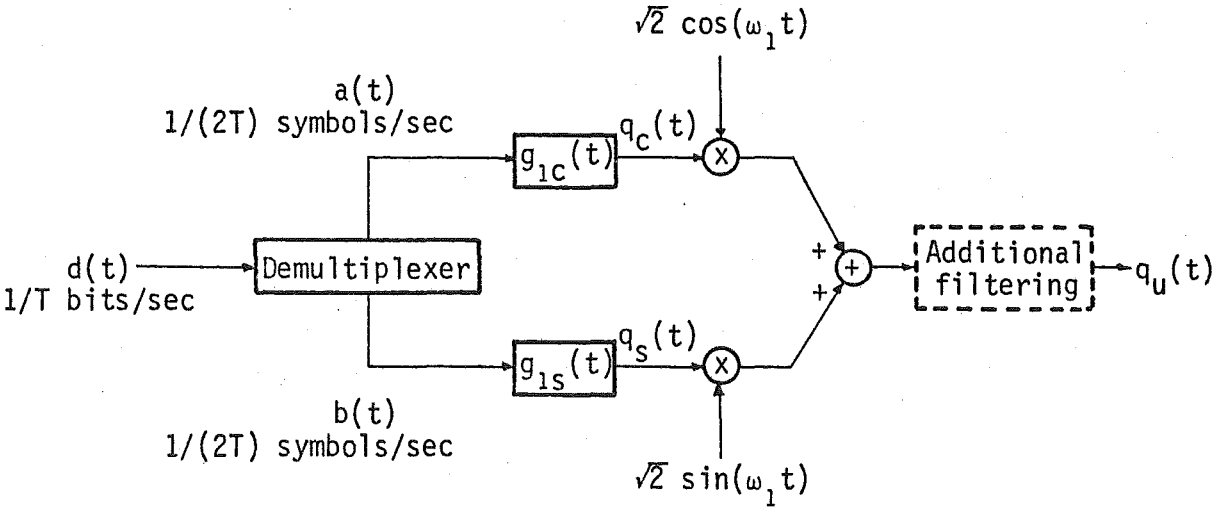


Fig. 5.2 Model for generating quadrature modulated signals.

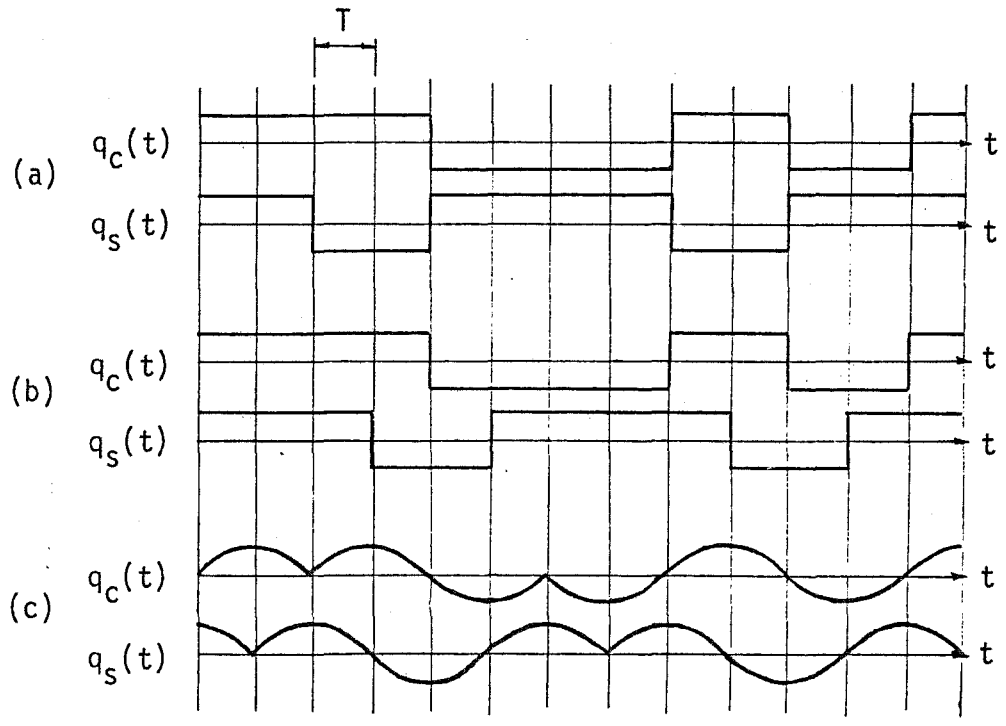


Fig. 5.3 Symbol shaping of data sequences for constant envelope signals; (a) QPSK, (b) OQPSK and (c) MSK.

$$A(t) = \{2(q_c^2(t) + q_s^2(t))\}^{\frac{1}{2}}, \tag{5.7}$$

and

$$\theta(t) = -\tan^{-1}\left[\frac{q_s(t)}{q_c(t)}\right]. \tag{5.8}$$

The signals with symbol shaping as shown in Fig. 5.3 will have constant envelopes, so no distortion would result if they were passed through a bandpass hardlimiter.

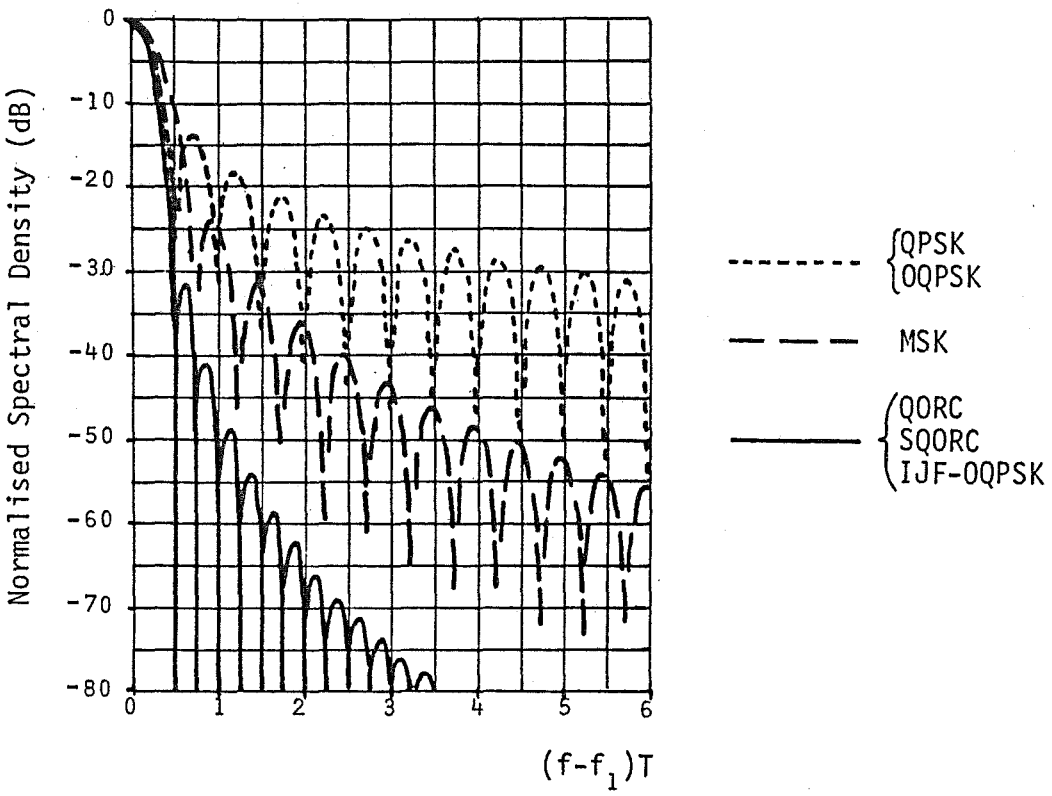


Fig. 5.4 Spectral densities of QPSK/OQPSK, MSK and QORC/SQORC/IJF-OQPSK.

The spectral efficiency of a signalling format must also be taken into consideration. The spectral densities of QPSK, OQPSK and MSK, with pulse shaping as in Fig. 5.3, are shown in Fig. 5.4. Each curve is obtained by taking the square of the Fourier transform of the relevant symbol shaping function [Pasupathy, 1979]. QPSK and OQPSK therefore have

identical spectral density curves. Fig. 5.4 indicates that signal power exists in sidelobes for each of the three signalling schemes. This sidelobe power will cause adjacent channel interference with any system in close frequency proximity. The interference can be reduced, thus allowing closer channel spacing, by using additional filtering as shown in Fig. 5.2. In practice this filtering would most likely be included in the baseband filtering functions $g_{1c}(t)$ and $g_{1s}(t)$.

Rhodes [1972] studied the effects of bandlimiting and hardlimiting on the spectral characteristics of QPSK, OQPSK and MSK. In (5.1) $\theta(t)$ is the instantaneous phase of $q_u(t)$ relative to that of the carrier. The instantaneous frequency, relative to the carrier, is given by

$$f_i(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}. \tag{5.9}$$

$f_i(t)$ will contain high frequency terms if $\theta(t)$ undergoes any sudden phase shifts. The frequency content of $f_i(t)$ is an indication of the degree to which a signal's spectral density is spread about its passband center; that is, the extent of spreading will increase as the high frequency power in $f_i(t)$ increases. In addition to $f_i(t)$, $A(t)$ will also have an effect on the spectral density. Envelope nulls or dips at points of high phase slope will suppress the spread of the signal power in the frequency domain that would otherwise occur if $A(t)$ was constant.

Prior to bandlimiting a constant envelope QPSK signal, $\theta(t)$ goes through phase shifts of 0, $\pm\pi/2$ and π radians. Bandlimiting QPSK will introduce envelope nulls at the points of π radian phase shifts. This is due to power in adjacent symbols being smeared into each other, so causing mutual cancellation at the mid-point of their overlap. The π radian phase shifts will still be present, but occur at points where $A(t)$ goes to zero. If the bandlimited QPSK signal is then hardlimited, the signal power at

points of high phase slope (π radian phase shifts) is restored to its level prior to bandlimiting, which in turn restores the spectral density of the signal back to its original form as shown in Fig. 5.4.

The use of additional filtering to significantly shape the signal spectrum following the nonlinearity is impractical in most cases due to the high carrier-to-signal bandwidth ratios involved. It is therefore necessary to use spectrally efficient modulation techniques which do not produce excessive out-of-band power when hardlimited.

In contrast to QPSK, Rhodes [1972] showed that for bandlimited OQPSK and MSK the effect of a hardlimiter caused a relatively small increase in out-of-band power. This is due to the avoidance of sudden π radian phase shifts. For OQPSK the relative time offset of the quadrature paths limits the sudden phase shifts to 0 or $\pm\pi/2$ radians. Bandlimiting smooths these phase transitions and causes a slight droop in the envelope in the region of the $\pm\pi/2$ radian shifts. Hardlimiting does not alter the smooth phase shifts and the restoration of the envelope to full amplitude introduces only a small amount of additional power in the spectral sidelobes.

MSK is a further step towards ensuring smooth phase transitions. This modulation method is also referred to as continuous-phase FSK (CPFSK) since, even without additional filtering, no sudden phase transitions occur. The phase of the MSK signal is continuously changing at $\pm\pi/(2T)$ rad/sec [Pasupathy, 1979], where the direction of the phase variation may change at symbol boundaries depending on the data sequence. Fig. 5.4 shows this results in MSK having less power in its spectral sidelobes compared with QPSK and OQPSK. Filtering MSK rounds off the corners where changes in the phase slope occur and introduces a slight droop in the amplitude of the envelope at these points. Consequently, hardlimiting introduces little additional sidelobe power to a bandlimited MSK signal.

In a more comprehensive study, Fang [1981] produced results which verified Rhodes' earlier work. Fang used a model which included both a HPA at the transmitter and a TWTA at the satellite transponder. He illustrated that increases in sidelobe power due to hardlimiting is at a minimum for signals with small envelope and smooth phase variations.

Since bandlimiting introduces envelope variations to QPSK, OQPSK and MSK, the advantage in using a signalling format which initially has a constant envelope is diminished. If high spectral efficiency is required it may be preferable to use a signalling format which does not have a constant envelope. Two such schemes are QORC and IJF-OQPSK. The transmitters for both schemes have the same form as shown in Fig. 5.2. Plots of their spectral densities are given in Fig. 5.4.

As the name implies, QORC uses symbol shaping pulses

$$g_{1c}(t) = g_{1s}(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{\pi t}{2T}\right)\right), & |t| \leq 2T \\ 0, & \text{elsewhere,} \end{cases} \tag{5.10}$$

where the symbol period is 2T seconds. A QORC signal suffers the same problem as bandlimited QPSK, in that its envelope goes through a null at points of π radian phase shifts. A large increase in sidelobe power therefore occurs when the QORC signal is hardlimited. This problem can be overcome by offsetting the quadrature related paths by half the symbol period, resulting in a signalling format called staggered QORC (SQORC).

A signalling format which is identical to SQORC is Feher's IJF-OQPSK. The distinguishing feature of Feher's system is that $g_{1c}(t)$ and $g_{1s}(t)$ operate as nonlinear processors or switches. Considering only $g_{1c}(t)$, since $g_{1s}(t)$ has identical logic, its basic processing logic is

as follows:

> When a(t) has 2 successive bits of the same polarity, $q_c(t)$ remains at a fixed d.c. level. If a(t) has a bit reversal, $q_c(t)$ will follow this reversal, taking a full 2T seconds for a smooth transition having the form of a sinusoidal function $\pm\cos(\pi t/(2T))$.

The advantage of IJF-OQPSK over SQORC is that signal shaping can be achieved by simply having d.c. and sinusoid sources which can be switched in or out.

In conjunction with spectral efficiency the BER performance of a system must also be considered. The distortion of a signal by a nonlinearity results in a performance degradation. This is usually considered in terms of the E_b/N_o degradation relative to the same system working over an ideal linear channel, with AWGN as the only perturbation. This degradation is considered in section 5.4 when VSB is compared with the modulation formats presented in this section.

5.3 HARDLIMITING VSB (AND OQPSK) SIGNALS

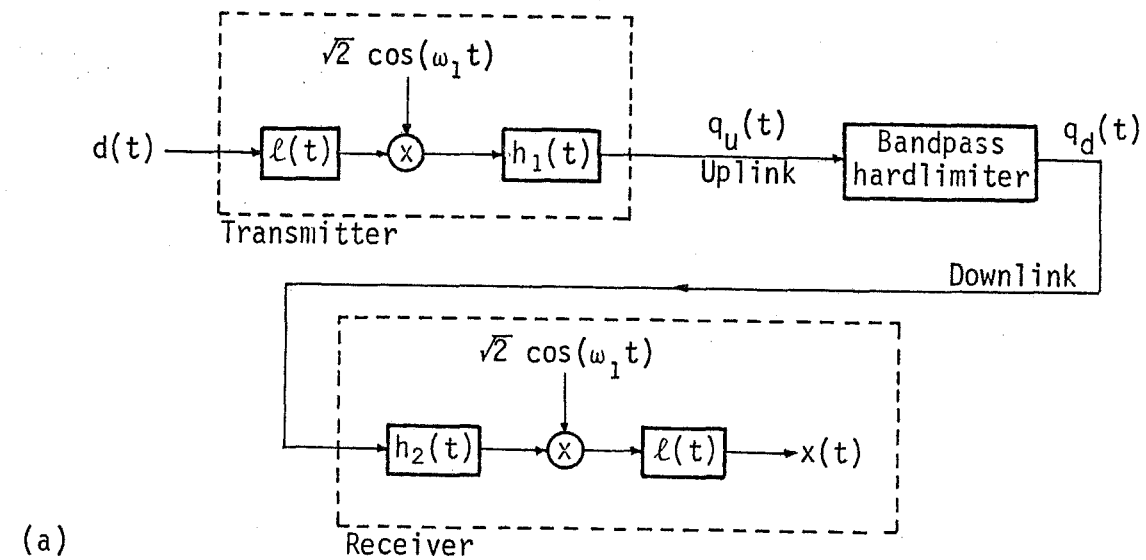
For the signalling formats presented in the previous section the symbol shaping functions all have a truncated time response. Consequently, this leads to the presence of sidelobes in the signals' spectral characteristics. No consideration appears to have been given to symbol shaping functions which are truncated in the frequency domain. This is the form of the VSB signal which is studied in this section.

Results of a computer simulation of VSB transmission over a nonlinear channel are presented. The similarity between VSB and OQPSK (see section 3.5) enables the results to be applied also to the latter signall-

ing format with cosine rolloff passband shaping. Also included are experimental results of hardlimiting a VSB signal which has an asymmetrical passband relative to its center frequency.

5.3.1 Simulation Models

Fig. 5.5(a) shows a block diagram of the simulated VSB system. The shaping of the VSB signal is by the transmitter and receiver filters with respective impulse responses $h_1(t)$ and $h_2(t)$. The frequency response of these filters is shown in Fig. 5.5(b). $h_1(t)$ and $h_2(t)$ have identical cosine rolloff frequency characteristics resulting in



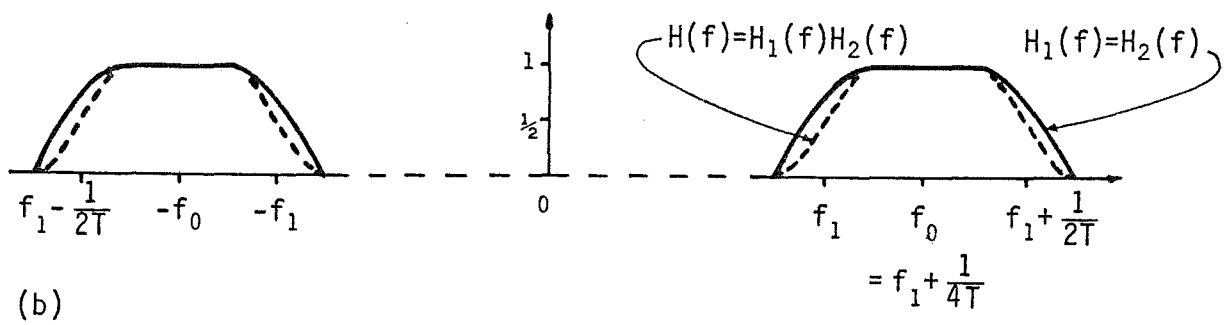


Fig. 5.5 (a) Simulated VSB system with nonlinear channel.
(b) Normalised frequency responses of identical transmitter and receiver filters (solid lines), and their cascade (dashed lines).

$$h(t) = h_1(t) * h_2(t) \tag{5.11}$$

having a raised-cosine frequency response with rolloff factor ρ. In the simulation, results for ρ equal to 1, 0.75 and 0.5 were obtained.

If no nonlinearity was present, a raised-cosine passband for h(t) would give optimum filtering with zero ISI. The presence of the bandpass hardlimiter will, however, distort the received VSB signal. To enable a comparison between VSB and other signalling formats, the effects which hardlimiting has on both the spectral density and BER performance were obtained. To avoid excessive computing time it was necessary to derive an equivalent baseband model to that shown in Fig. 5.5(a). Also, since actual plots of P_b versus E_b/N_o would require very long computing times, eye-pattern diagrams were used to give an indication of the expected E_b/N_o degradation. The validity of this is considered in section 5.3.2 when the results are analysed.

Using complex representation (see section 3.2.1), $q_u(t)$ can be defined as

$$q_u(t) = \sqrt{2}\ \mathrm{Re}\left[\left\{d(t) * \frac{h_{1c}(t)}{2} + j\ d(t) * \frac{h_{1s}(t)}{2}\right\} e^{j\omega_1 t}\right], \tag{5.12}$$

where $h_{1c}(t)$ and $h_{1s}(t)$ make up the complex envelope of $h_1(t)$. The output of the bandpass hardlimiter will therefore be

$$q_d(t) = \frac{\sqrt{2}\ \mathrm{Re}\left[\left\{d(t) * \frac{h_{1c}(t)}{2} + j\ d(t) * \frac{h_{1s}(t)}{2}\right\} e^{j\omega_1 t}\right]}{\left|d(t) * \frac{h_{1c}(t)}{2} + j\ d(t) * \frac{h_{1s}(t)}{2}\right|}. \tag{5.13}$$

The denominator in (5.13) removes the envelope variations in $q_u(t)$, but does not alter its phase. Letting

$$q_c(t) = d(t) * \frac{h_{1c}(t)}{2} \tag{5.14a}$$

and

$$q_s(t) = d(t) * \frac{h_{1s}(t)}{2}, \tag{5.14b}$$

then from (5.13) the complex envelope of the hardlimited VSB signal is

$$\sqrt{2}\ \tilde{q}_d(t) = \frac{\sqrt{2}\ \{q_c(t) + j\ q_s(t)\}}{|q_c(t) + j\ q_s(t)|}\ . \tag{5.15}$$

Haykin [1978] indicates a method by which the spectral density of $q_d(t)$ can be estimated using $\tilde{q}_d(t)$. Let $q_d^\tau(t)$ and $\tilde{q}_d^\tau(t)$ define truncated time samples of, respectively, $q_d(t)$ and $\tilde{q}_d(t)$, where τ is the length of the truncation. $q_d^\tau(t)$ will have an amplitude spectrum given by

$$|Q_d^\tau(f)| = |F[q_d^\tau(t)]|. \tag{5.16}$$

Using (5.16), the spectral density of $q_d(t)$ is

$$S_d(f) = \lim_{\tau\to\infty} \frac{1}{\tau} E[|Q_d^\tau(f)|^2]. \tag{5.17}$$

If $|\tilde{Q}_d^\tau(f)|$ is the amplitude spectrum of $\tilde{q}_d^\tau(t)$, then its shape about 0 Hz will be identical to the shape of $|Q_d^\tau(f)|$ about f_1 Hz. The same relationship will therefore exist between $S_d(f)$ and $\tilde{S}_d(f)$, the spectral density of the complex envelope, where

$$\tilde{S}_d(f) = \lim_{\tau\to\infty} \frac{1}{\tau} E[|\tilde{Q}_d^\tau(f)|^2]. \tag{5.18}$$

The baseband model in Fig. 5.6(a) follows from (5.14), (5.15) and (5.18). In the computer simulation τ did not go to infinity, so $\tilde{S}_d(f)$ (hence

$S_d(f)$) was found by averaging the results of many computer runs with different random input data.

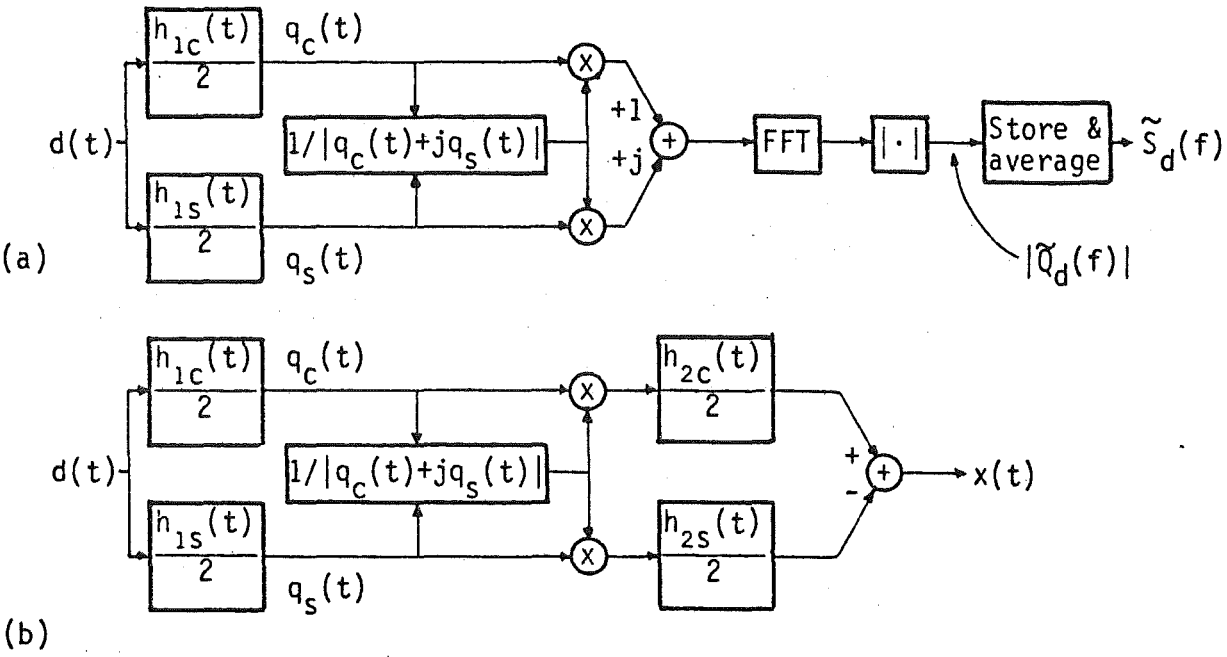


Fig. 5.6 Equivalent baseband models of hardlimited VSB system for obtaining (a) spectral density and (b) eye-pattern plots.

The recovered baseband signal in the hardlimited VSB system will be

$$x(t) = \tilde{q}_{d_c}(t) * \frac{h_{2c}(t)}{2} - \tilde{q}_{d_c}(t) * \frac{h_{2s}(t)}{2}, \tag{5.19a}$$

where

$$\tilde{q}_{d_c}(t) = \mathrm{Re}[\tilde{q}_d(t)], \tag{5.19b}$$

$$\tilde{q}_{d_s}(t) = \mathrm{Im}[\tilde{q}_d(t)], \tag{5.19c}$$

and $h_{2c}(t)$ and $h_{2s}(t)$ make up the complex envelope of $h_2(t)$. Fig. 5.6(b), the equivalent baseband model for deriving x(t), follows from (5.14), (5.15) and (5.19).

The similarity between VSB and OQPSK was demonstrated in section 3.5, where it was shown that with simple encoding/decoding, VSB modulation/

demodulation could be implemented using OQPSK circuitry. For random data and a given transmitter passband shaping (cosine rolloff in this case), the spectral density functions for VSB and OQPSK will therefore be identical, both before and after hardlimiting. The baseband signals for OQPSK can be derived by modifying the model in Fig. 5.6(b). The eye-pattern is the result of interest and not the correct data reception in terms of bit polarity. The hardlimited VSB signal (equ. (5.13)) can therefore be used to also define a hardlimited OQPSK signal. Filtering $q_d(t)$ by $h_2(t)$, then quadrature demodulating by $\sqrt{2}\cos(\omega_0 t)$ and $\sqrt{2}\sin(\omega_0 t)$, where $f_0 = f_1 + 1/(4T)$, will result in two baseband signals, respectively,

$$y(t) = \frac{1}{2}\Big[\{\tilde{q}_{d_c}(t) * h_{2c}(t) - \tilde{q}_{d_s}(t) * h_{2s}(t)\}\cos\left(\frac{\pi t}{2T}\right) + \{\tilde{q}_{d_c}(t) * h_{2s}(t) + \tilde{q}_{d_s}(t) * h_{2c}(t)\}\sin\left(\frac{\pi t}{2T}\right)\Big] \tag{5.20a}$$

and

$$z(t) = \frac{1}{2}\Big[\{\tilde{q}_{d_c}(t) * h_{2c}(t) - \tilde{q}_{d_s}(t) * h_{2s}(t)\}\sin\left(\frac{\pi t}{2T}\right) - \{\tilde{q}_{d_c}(t) * h_{2s}(t) + \tilde{q}_{d_s}(t) * h_{2c}(t)\}\cos\left(\frac{\pi t}{2T}\right)\Big] \tag{5.20b}$$

Fig. 5.7 shows the OQPSK simulation model used to derive y(t) and z(t).

Some insight into the effect of including the bandpass hardlimiter in the channel can be obtained by observing the uplink signal's carrier envelope A(t) and phase $\theta(t)$. These two terms are given in (5.7) and (5.8), and can be derived from the transmitter stage of the simulation models in either Fig. 5.6 or Fig. 5.7.

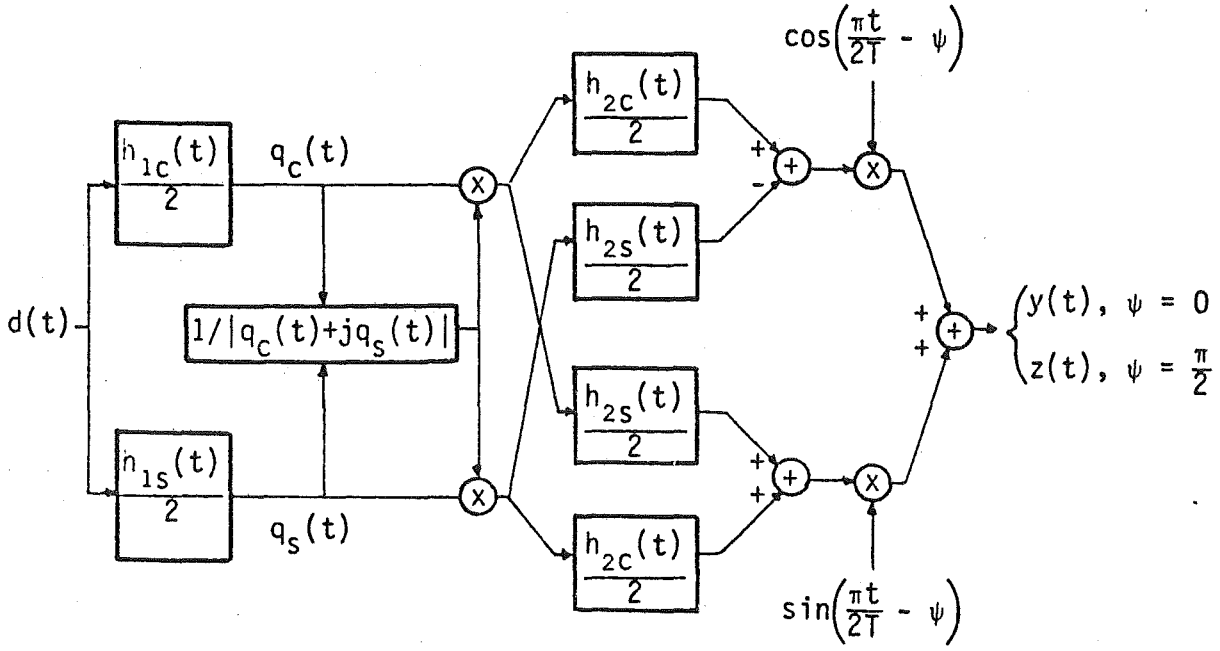


Fig. 5.7 Equivalent baseband model of hardlimited OQPSK for obtaining eye-patterns.

5.3.2 Simulation Results and Analysis

Plots of the spectral densities of the VSB/OQPSK signals before and after hardlimiting are shown in Fig. 5.8. The plots indicate that the power introduced outside the transmitted signals' passbands due to hardlimiting is relatively independent of ρ. The reason for this can be obtained from the associated plots of carrier envelope and phase prior to hardlimiting, with the same data sequence used for each plot. The variations in $\theta(t)$, hence $f_i(t)$, for the different ρ values are very similar. Prior to hardlimiting, the dips in the envelopes ensure the signals' power content at points of high instantaneous frequency is suppressed. An example of such a point is shown in Fig. 5.8. The depth of the envelope dips increase with the narrowing of the VSB passband. Following hardlimiting, $\theta(t)$ remains unchanged but $A(t)$ becomes constant, so the power at points of high instantaneous frequency is no longer suppressed. Power is therefore introduced outside the transmitted signals' passbands. With the

envelope constant, the spectral density of a signal will be determined by its carrier phase, hence with $\theta(t)$ relatively independent of ρ the out-of-band power introduced by hardlimiting is similar for the examples given.

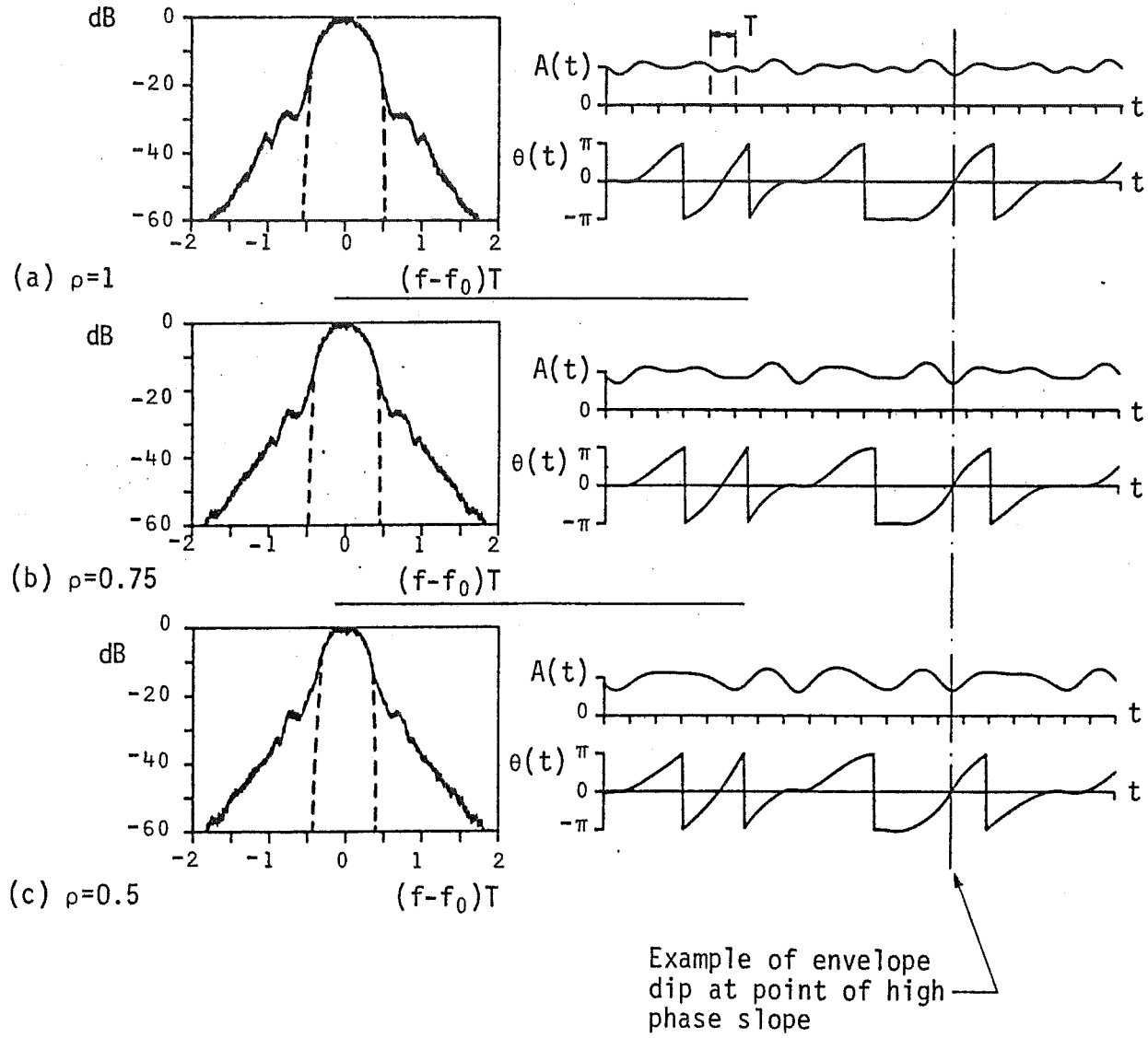


Fig. 5.8 Normalised spectral densities of VSB/OQPSK signals before (dashed lines) and after (solid lines) hard-limiting, and typical carrier envelope and phase plots. Transmitted signal passbands are cosine rolloff with indicated ρ.

The eye-patterns for the VSB system with and without hardlimiting are shown in Fig. 5.9. Hardlimiting is seen to cause increasing eye-pattern degradation as ρ decreases. The plots of carrier envelopes in

Fig. 5.8 show the reason for this. Hardlimiting introduces AM-AM distortion, with envelope variations in $q_u(t)$ being completely removed. Distortion introduced by hardlimiting will be at a minimum for signals with small envelope variations, therefore hardlimiting causes the least degradation to the VSB signal with a passband rolloff of $\rho = 1$.

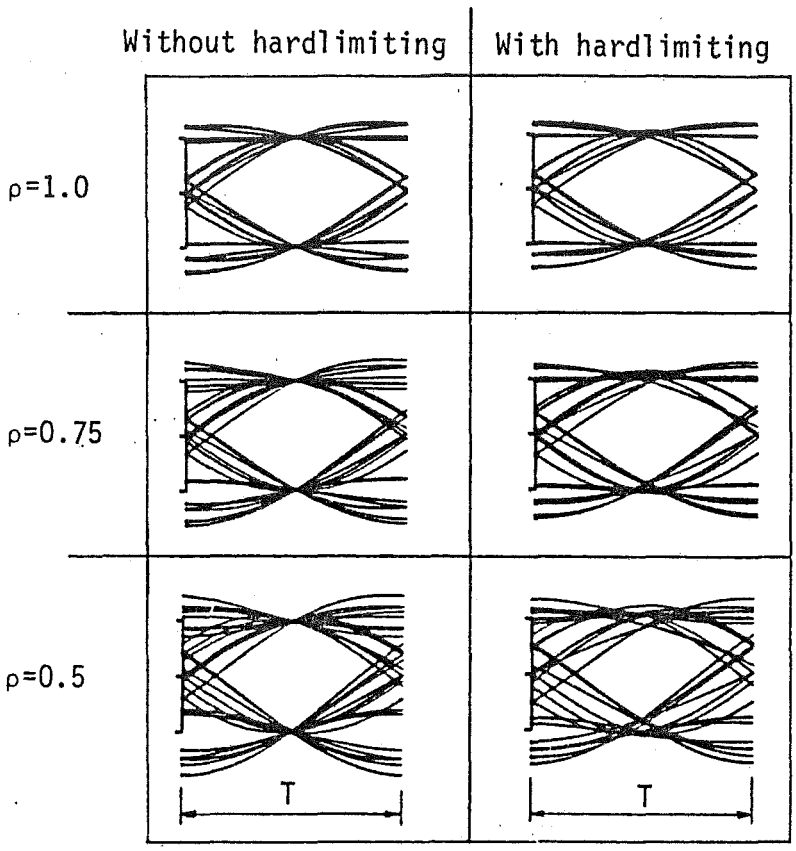


Fig. 5.9 Eye-patterns for VSB with and without hardlimiting.

Fig. 5.10 shows eye-pattern plots for OQPSK with and without hardlimiting. For a given value of ρ the degradation due to hardlimiting, in terms of eye-pattern closer at the optimum sampling point, is identical to VSB.

The expected E_b/N_o degradation caused by hardlimiting can be determined from the eye-pattern plots if two assumptions are made; (a) the only interference other than hardlimiting is an AWGN source which is introduced after the nonlinearity, and (b) the downlink E_b/N_o is large enough so

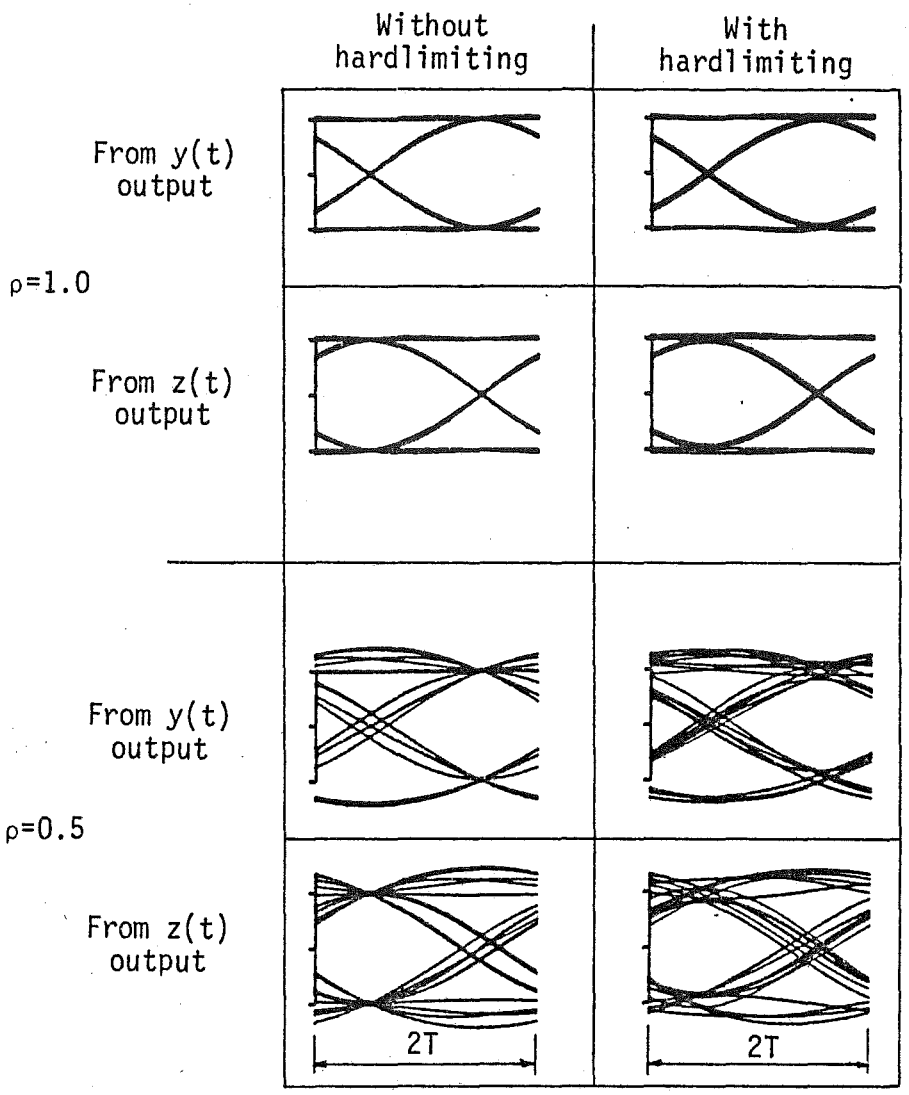


Fig. 5.10 Eye-pattern plots for OQPSK with and without hardlimiting.

that the BER performance is determined by the minimum eye-pattern opening. Assumption (a) is one that has been made in several studies where the channel is a satellite link. Kennedy and Shimbo [1981] indicate that it is an approximation which is justified in most practical applications. If the downlink E_b/N_o is a few decibels worse than the uplink, as often exists, the contribution of the uplink noise to overall BER is negligible. Gronemeyer and McBride [1976] arrive at a similar conclusion where, in their simulation model for MSK and OQPSK, uplink E_b/N_o is maintained at a relatively high level. Assumption (b) makes use of the AWGN having a

linear effect on the recovered baseband signal. If the required BER for the VSB system is small, less than 10^{-5} say, the errors which do occur will be determined mainly by the sampled baseband signal levels which are closest to the decoder's decision threshold. These levels correspond to the minimum openings in the eye-patterns shown in Fig. 5.9 and Fig. 5.10. The approximate E_b/N_o degradation can therefore be calculated from the relative eye-pattern openings of the systems with and without hardlimiting. These values are given in Table 5.1.

TABLE 5.1

APPROXIMATE E_b/N_o DEGRADATIONS WHEN VSB/OQPSK SIGNALS WITH INDICATED PASSBAND ROLLOFF FACTORS UNDERGO HARDLIMITING

ρ	E_b/N_o Degradation
1.0	0.5 dB
0.75	0.9 dB
0.5	1.7 dB

Adjacent channel interference also contributes to performance degradation. If two or more channels are frequency multiplexed in the same transmission medium, it is desirable, in terms of efficient spectral usage, to position the channels in as close frequency proximity as possible. How close adjacent channels can be placed depends on the extent to which the signal power in one channel interferes with the other; i.e., on the amount of overlap in their spectral densities. With the fast rolloff characteristic of the cosine filter, for the uplink path adjacent channel interference will be negligible compared with the hardlimited signal where power is spread out in the frequency domain. No simulation results of the effect of adjacent channel interference were obtained. However, in section 5.4, where simulation results are compared with the

systems introduced in section 5.2, it is shown that the rate of rolloff of the spectral density of hardlimited VSB is on a par with the alternative systems. A similar degradation due to adjacent channel interference could therefore be expected.

5.3.3 Experimental Results From Hardlimiting VSB

As discussed in section 3.6, a 9600 bits/sec VSB system was designed and constructed using analogue passband shaping filters positioned to enable in-band generation. It was a simple task to include a hardlimiter, as per Fig. 5.5(a), in the experimental system. The results obtained show the interesting effect of hardlimiting a signal with asymmetrical passband shaping about its center frequency.

The time waveforms of the experimental VSB signal at the input to the receiver, with and without hardlimiting in the channel, are shown in Fig. 5.11 along with the associated eye-pattern plots. The eye-pattern degradation due to hardlimiting is evident.

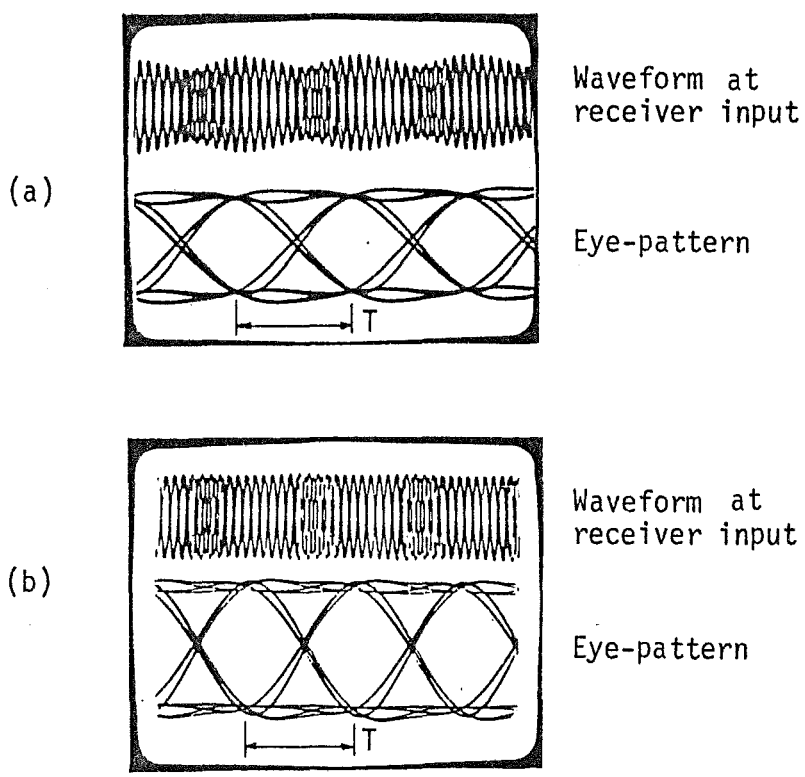


Fig. 5.11 Oscilloscope tracings of received VSB waveforms and associated eye-patterns; (a) without and (b) with hardlimiting.

The use of analogue bandpass filters for VSB signal shaping resulted in asymmetry in the passband about its center frequency (60 kHz in this case). The effect hardlimiting has on its spectral density is shown in Fig. 5.12. The spectral shaping of the VSB signal at the transmitter output has slower rolloff characteristic above 60 kHz compared with the rolloff rate below 60 kHz. With hardlimiting, it is the region below 60 kHz which suffers the greater introduction of out-of-band power as the spectrum of the hardlimited signal tends towards being symmetrical about its center frequency.

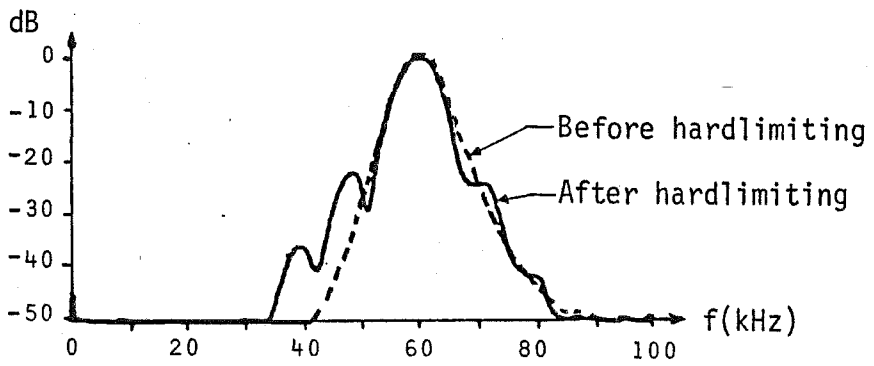


Fig. 5.12 Spectral density plots of experimental VSB signal before and after hardlimiting. Recordings taken using a Nicolet Scientific Corp. 444A Mini-Ubiquitous FFT Analyser.

5.4 COMPARISON OF HARDLIMITED VSB WITH ALTERNATING SYSTEMS

Using the results from section 5.3.2, VSB with raised-cosine rolloff passband shaping is compared with the alternative systems introduced in section 5.2. Care must be exercised in accepting the relative figures per se, since the models used for the various systems in this comparison are often more complex than the nonlinear VSB model of Fig. 5.5(a). The important differences will be noted.

Since decreasing the VSB passband rolloff gives no significant advantage in terms of spectral efficiency of the hardlimited signal, the

results for the passband where $\rho = 1$ will be used. For this case the E_b/N_o degradation due to hardlimiting is at a minimum.

Some basis of comparison is required. In line with the simulation results this will be the E_b/N_o degradation and the spectral efficiency at the output of the hardlimiter. A measure of the spectral efficiency will be taken as the relative levels of the spectral densities at $f_0 \pm 1/T$ and $f_0 \pm 1.5/T$ (i.e., $S_d(f_0 \pm 1/T)$ and $S_d(f_0 \pm 1.5/T)$) with respect to the spectral density at the passband center frequency f_0.

From Table 5.1, hardlimited VSB, when $\rho = 1$, suffers an E_b/N_o degradation of approximately 0.5 dB, and from Fig. 5.8, $S_d(f_0 \pm 1/T) = -37$ dB and $S_d(f_0 \pm 1.5/T) = -56$ dB.

For QPSK, OQPSK and MSK signals with constant envelopes, no E_b/N_o degradation is suffered due to hardlimiting. The spectral densities of the signals after hardlimiting will remain as shown in Fig. 5.4. For QPSK and OQPSK, $S_d(f_0 \pm 1/T) = -17$ dB and $S_d(f_0 \pm 1.5/T) = -21$ dB. The spectral density plots actually go to $-\infty$ dB at these points, but this would give a misleading indication of rolloff. The values given here, and for the following cases, assumes a smooth line has been drawn through the peaks of the spectral lobes. For MSK, $S_d(f_0 \pm 1/T) = -23$ dB and $S_d(f_0 \pm 1.5/T) = -31$ dB. Though these constant envelope systems suffer no E_b/N_o degradation, their spectral efficiency is poor and additional filtering is required to increase this efficiency. This filtering will introduce carrier envelope variations, hence AM-AM distortion by the hardlimiter, resulting in an E_b/N_o degradation in performance.

Mathwich, et al. [1974] studied the effect of filtering followed by hardlimiting of MSK. The nonlinear model used in the study was very similar to the simulated VSB system. An ideal MSK receiver was used, consisting of symbol weighting followed by an integrate-and-dump stage for

each quadrature channel. Several different filtering functions were tested in the study. The result of interest is when the filtered and hardlimited MSK signal had a spectral density characteristic similar to the VSB system. This occurred for a Gaussian filter at the transmitter with a -3 dB bandwidth of 0.71/T Hz. For this case, $S_d(f_0 \pm 1/T) = -37$ dB and $S_d(f_0 \pm 1.5/T) = -50$ dB. With a filter also at the receiver to give an overall -3 dB bandwidth of 0.64/T Hz, the experimental E_b/N_o degradation was approximately 2 dB. However, the theoretical degradation for this system was given as only approximately 0.3 dB. This latter figure gives a more fair comparison with the VSB model which was also ideal. From these figures it appears as though there is little difference in the performance of VSB with cosine rolloff shaping filters and MSK with additional filtering at the transmitter and receiver.

Fang [1981] studied the combined effects of nonlinear distortion, fading and cochannel interference on QPSK, OQPSK and MSK. The model he used was more representative of an actual satellite link, making it difficult to compare with the results for the simple nonlinear VSB model. Fang's results possibly give a better indication of what could be expected from a more comprehensive VSB model or an actual VSB system. With "special pulse shaping", OQPSK had spectral density rolloff given by $S_d(f_0 \pm 1/T) = -34$ dB and $S_d(f_0 \pm 1.5/T) = -40$ dB, with an E_b/N_o degradation of 1.4 dB at $P_b = 10^{-5}$. For MSK, "special pulse shaping" gave $S_d(f_0 \pm 1/T) = -34$ dB and $S_d(f_0 \pm 1.5/T) = -41$ dB, with an E_b/N_o degradation at $P_b = 10^{-5}$ of 1.8 dB.

The study of SQORC modulation by Austin and Chang [1981] used a model that consisted of a phase equalised channel filter prior to a TWTA, with AWGN on the downlink. The spectral density of the SQORC signal at the output of the TWTA was much the same as for the VSB system, with $S_d(f_0 \pm 1/T) = -42$ dB and $S_d(f_0 \pm 1.5/T) = -57$ dB. The overall E_b/N_o

degradation was 2.1 dB at $P_b = 10^{-5}$. Again it is difficult to compare the results with those for VSB, since the SQORC model included modulator imbalance and a channel filter with nonlinear group delay characteristics.

Austin and Chang [1981] also obtained results for OQPSK and MSK operating over the same channel as for the SQORC system. For the OQPSK system, $S_d(f_0 \pm 1/T) = -37$ dB and $S_d(f_0 \pm 1.5/T) = -51$ dB, with an E_b/N_o degradation of 1.7 dB at $P_b = 10^{-5}$. For the MSK system, $S_d(f_0 \pm 1/T) = -28$ dB and $S_d(f_0 \pm 1.5/T) = -40$ dB with an E_b/N_o degradation of 0.9 dB at $P_b = 10^{-5}$.

As could be expected, IJF-OQPSK has identical spectral characteristics to SQORC. One of the IJF-OQPSK system models evaluated by Le-Ngoc, et al. [1982] was similar to the VSB model, in that it consisted of a hard-limited single-channel system with a bandlimiting receiver filter and AWGN on the downlink. The minimum E_b/N_o degradation for this system was 1 dB at $P_b = 10^{-4}$.

With the relatively simple nonlinear VSB model, E_b/N_o degradation was similar to the theoretical degradation of filtered MSK which Mathwich, et al. [1974] derived. The system perturbations in both cases were a hard-limiter followed by downlink AWGN. The more comprehensive experimental and simulation results for the various systems indicated the E_b/N_o degradation was generally in the order of 1 to 2 dB. It is reasonable to expect a similar degradation to occur if a more comprehensive VSB model was used.

A simplified breakdown of the results for the various systems used in the comparison is given in Table 5.2.

TABLE 5.2

COMPARISON OF SYSTEMS OPERATING OVER NONLINEAR CHANNELS

System	Channel Distortion	$S_d(f_0 \pm 1/T)$	$S_d(f_0 \pm 1.5/T)$	E_b/N_o degradation
VSB (cosine filters)	hardlimiter	-37 dB	-56 dB	0.5 dB
QPSK/OQPSK (constant envelope)	hardlimiter	-17 dB	-21 dB	0 dB
MSK (constant envelope)	hardlimiter	-23 dB	-31 dB	0 dB
MSK (filtered)	hardlimiter	-37 dB	-50 dB	2 dB (experimental) 0.3 dB (theoretical)
OQPSK (special pulse shaping)	HPA, TWTA, cochannel interference	-34 dB	-40 dB	1.4 dB
MSK (special pulse shaping)	HPA, TWTA, cochannel interference	-34 dB	-41 dB	1.8 dB
SQORC (filtered)	modulator imbalance, group delay, TWTA	-42 dB	-57 dB	2.1 dB
OQPSK (filtered)	modulator imbalance, group delay, TWTA	-37 dB	-51 dB	1.7 dB
MSK (filtered)	modulator imbalance, group delay, TWTA	-28 dB	-40 dB	0.9 dB
IJF-OQPSK	hardlimiter	-42 dB	-57 dB	1 dB

CHAPTER 6

QUADRATURE VESTIGIAL SIDEBAND (QVSB) DATA TRANSMISSION

6.1 INTRODUCTION

The use of VSB modulation for the transmission of data signals offers an attractive compromise between DSB and SSB methods. Its spectral efficiency approaches that of SSB, but filtering requirements do not become too severe. Another alternative for increasing spectral efficiency is the use of quadrature modulation. This technique is, in general, restricted to DSB systems where two signals modulate quadrature related carriers to produce two non-interfering signals.

In this chapter a new technique which combines quadrature modulation and VSB techniques is introduced and analysed. QVSB is chosen as a name which aptly describes the system's modulation format. A QVSB signal basically consists of two independent data streams which are VSB modulated into the same passband on separate quadrature related carriers. Such a scheme is shown to introduce ISI between the two VSB signals in a similar manner to PRS systems. Methods of "unwinding" this ISI are derived, ranging from a simple precoding technique which allows bit-by-bit decoding, to the more complex probabilistic decoding techniques which give improved BER performance.

6.2 QVSB MODEL

The QVSB system is described by the model shown in Fig. 6.1. Data sequences

$$\underline{a} = (a_0, a_1, a_2, \ldots, a_N) \qquad (6.1)$$

and

$$\underline{b} = (b_0, b_1, b_2, \ldots, b_N) \tag{6.2}$$

each consist of independent unit impulses spaced at T_c second intervals, where $a_k = \pm 1$ and $b_k = \pm 1$. The two sequences could be considered as being derived by splitting an initial sequence with information rate 1/T bits/sec, where $T_c = 2T$. $\underline{a}$ and $\underline{b}$ can be alternatively described by the time functions, respectively,

$$a(t) = \sum_{k=0}^{N} a_k \, \delta(t - kT_c) \tag{6.3}$$

and

$$b(t) = \sum_{k=0}^{N} b_k \, \delta(t - kT_c). \tag{6.4}$$

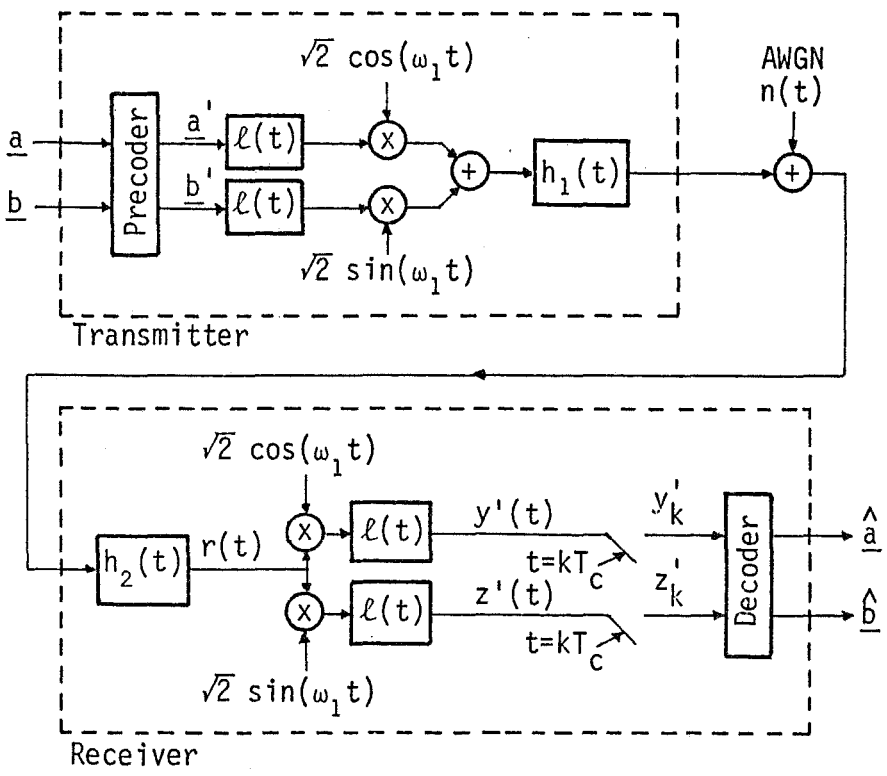


Fig. 6.1 Model of QVSB system.

The outputs of the precoder stage will have the form

$$a'(t) = f_1(a(t), b(t))$$

$$= \sum_{k=0}^{N} a'_k \, \delta(t - kT_c) \tag{6.5}$$

and

$$b'(t) = f_2(a(t), b(t))$$

$$= \sum_{k=0}^{N} b'_k \, \delta(t - kT_c). \tag{6.6}$$

The inclusion of this stage is mainly for a technique devised by Lender [1963, 1964, 1966] known as precoding. Lender applied this technique to PRS systems to enable bit-by-bit decoding of the received signal. Precoding of QVSB is analysed in section 6.4.

At the transmitter, the lowpass filters with impulse response $\ell(t)$ remove high frequency terms in $a'(t)$ and $b'(t)$ (due to $\delta(t - kT_c)$) prior to modulation. Similarly, the lowpass filters at the receiver remove sum frequencies at the demodulator outputs prior to decoding.

It will be assumed that QVSB passband shaping is determined by the bandpass filters with impulse responses $h_1(t)$ and $h_2(t)$. Let

$$h(t) = h_1(t) * h_2(t) \tag{6.7}$$

and

$$H(f) = F[h(t)]. \tag{6.8}$$

$H(f)$ is chosen to have a raised-cosine passband, initially with a rolloff factor $\rho = 1$ as shown in Fig. 6.2.

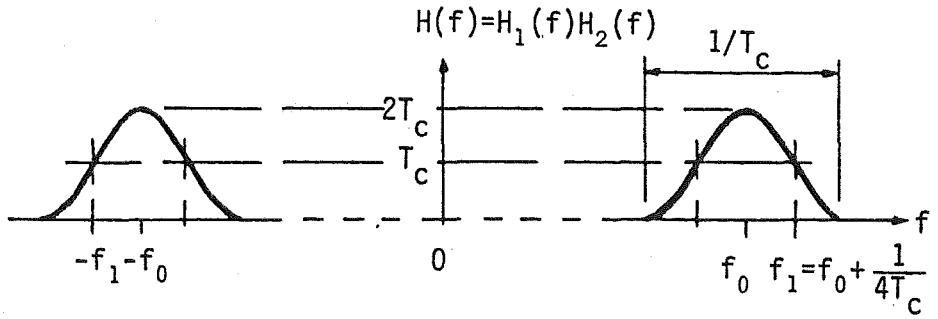


Fig. 6.2 QVSB passband when $\rho = 1$.

Modulation of the lowpass filtered a'(t) and b'(t) signals on to quadrature related carriers $\sqrt{2}\cos(\omega_1 t)$ and $\sqrt{2}\sin(\omega_1 t)$, respectively, and then filtering by h(t), results in r(t) consisting of two VSB signals sharing the same passband. Mutual ISI between these two VSB signals must exist since any VSB signal has an inherent quadrature component. This ISI is carried through to the recovered baseband signals $\underline{y}$ and $\underline{z}$. Non-linear processing by the precoder and/or decoder stages is required to resolve the received information and obtain estimates $\hat{\underline{a}}$ and $\hat{\underline{b}}$ of $\underline{a}$ and $\underline{b}$, respectively.

6.3 ANALYSIS OF QVSB MODEL

The model in Fig. 6.1 is analysed to derive the ISI between the two VSB signals on quadrature related carriers.

With the system passband having raised-cosine shaping, its time response can be defined as

$$h(t) = 2g(t)\cos(\omega_0 t)$$

$$= 2g(t)\cos\left(\frac{\pi t}{2T_c}\right)\cos(\omega_1 t) + 2g(t)\sin\left(\frac{\pi t}{2T_c}\right)\sin(\omega_1 t), \quad (6.9)$$

where

$$g(t) = \frac{\text{sinc}\left(\frac{t}{2T_c}\right) \cos\left(\frac{\pi\rho t}{2T_c}\right)}{1 - \rho^2 t^2/T_c^2} . \qquad (6.10)$$

Initially the rolloff factor ρ is assumed to be unity. Decreasing ρ is considered in section 6.5.4.

To maximise SNR at the input to the decoder,

$$h_2(t) = h_1(-t)$$

$$= h_1(t) \qquad (6.11)$$

since h(t) is symmetrical about t = 0.

The precoder stage in Fig. 6.1 is only included in some specific forms of the QVSB model and unless otherwise stated it will be assumed that $\underline{a}' = \underline{a}$ and $\underline{b}' = \underline{b}$; i.e., there is no precoding. The output of $h_2(t)$ is therefore given by

$$r(t) = \sqrt{2}\ \text{Re}\left[\left\{(a(t) - j\ b(t)) * \left(g(t) \cos\left(\frac{\pi t}{2T_c}\right)\right.\right.\right.$$

$$\left.\left.\left. - j\ g(t) \sin\left(\frac{\pi t}{2T_c}\right)\right)\right\} e^{j\omega_1 t}\right] + \beta(t), \qquad (6.12)$$

where $\beta(t)$ is the filtered AWGN;

$$\beta(t) = n(t) * h_2(t). \qquad (6.13)$$

$\beta(t)$ can be defined as [Haykin, 1978]

$$\beta(t) = n_c(t) \cos(\omega_0 t) - n_s(t) \sin(\omega_0 t), \qquad (6.14)$$

where $n_c(t)$ and $n_s(t)$ are independent Gaussian baseband variables, both with the same variance as $\beta(t)$.

After demodulating r(t),

$$y'(t) = y(t) + n_y(t) \tag{6.15}$$

and

$$z'(t) = z(t) + n_z(t), \tag{6.16}$$

where

$$y(t) = a(t) * \left\{g(t) \cos\left(\frac{\pi t}{2T_c}\right)\right\} - b(t) * \left\{g(t) \sin\left(\frac{\pi t}{2T_c}\right)\right\} \tag{6.17}$$

and

$$z(t) = b(t) * \left\{g(t) \cos\left(\frac{\pi t}{2T_c}\right)\right\} + a(t) * \left\{g(t) \sin\left(\frac{\pi t}{2T_c}\right)\right\} \tag{6.18}$$

are the signal terms, with the Gaussian noise terms given by

$$n_y(t) = \frac{1}{\sqrt{2}} n_c(t) \cos\left(\frac{\pi t}{2T_c}\right) + \frac{1}{\sqrt{2}} n_s(t) \sin\left(\frac{\pi t}{2T_c}\right) \tag{6.19}$$

and

$$n_z(t) = \frac{1}{\sqrt{2}} n_c(t) \sin\left(\frac{\pi t}{2T_c}\right) - \frac{1}{\sqrt{2}} n_s(t) \cos\left(\frac{\pi t}{2T_c}\right). \tag{6.20}$$

Letting

$$g_c(t) = g(t) \cos\left(\frac{\pi t}{2T_c}\right) \tag{6.21}$$

and

$$g_s(t) = g(t) \sin\left(\frac{\pi t}{2T_c}\right), \tag{6.22}$$

then in (6.17) and (6.18), $g_c(t)$ could be thought of as the desired response to a data bit and $g_s(t)$ as the response which introduces ISI. These two terms are shown in Fig. 6.3.

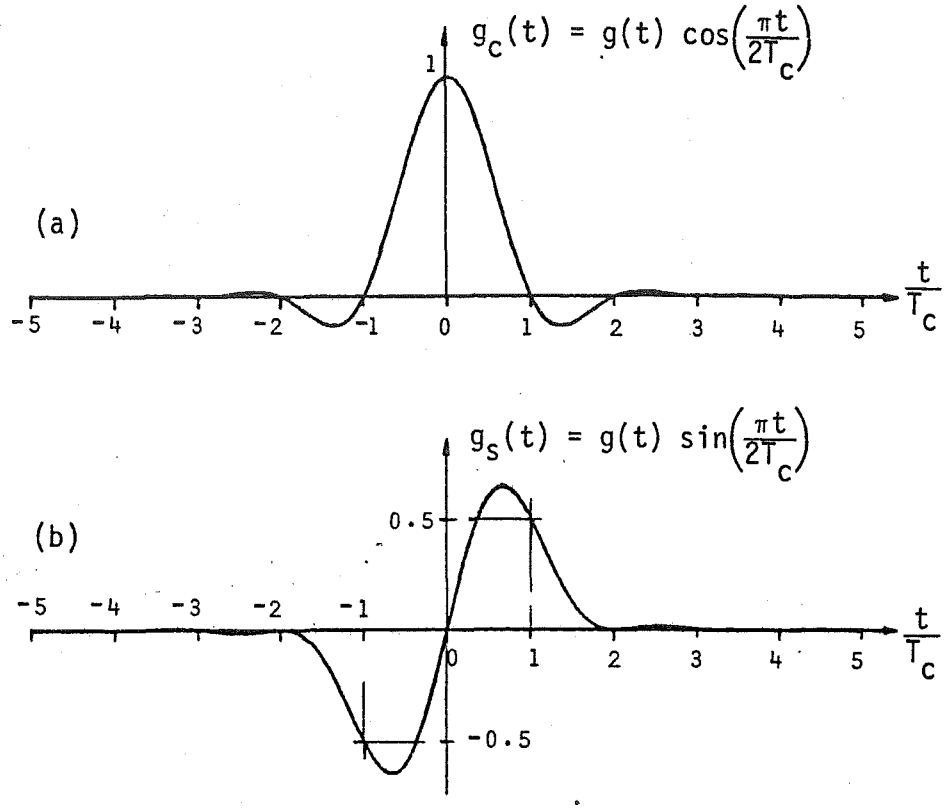


Fig. 6.3 (a) Desired and (b) interference data responses in QVSB system when $\rho = 1$.

The sampled inputs to the decoder will have the form

$$y'_k = y'(kT_c)$$

$$= y_k + n_{y_k} \tag{6.23}$$

and

$$z'_k = z'(kT_c)$$

$$= z_k + n_{z_k}. \tag{6.24}$$

Using (6.17),

$$y_k = \int_{-\infty}^{\infty} y(t)\delta(t - kT_c)dt$$

$$= \int_{-\infty}^{\infty} \left[\left\{ \sum_{i=0}^{N} a_i \, \delta(t - iT_c) \right\} * g_c(t) \right.$$

$$\left. - \left\{ \sum_{i=0}^{N} b_i \, \delta(t - iT_c) \right\} * g_s(t) \right] \delta(t - kT_c)dt$$

$$= \int_{-\infty}^{\infty} \left[\sum_{i=0}^{N} \left\{ a_i \, g_c(t - iT_c) - b_i \, g_s(t - iT_c) \right\} \right] \delta(t - kT_c)dt$$

$$= \sum_{i=0}^{N} \left\{ a_i \, g_c((k-i)T_c) - b_i \, g_s((k-i)T_c) \right\}. \qquad (6.25)$$

From Fig. 6.3 it can be seen that $g_c((k-i)T_c)$ is non-zero only when $(k-i) = 0$, and $g_s((k-i)T_c)$ is non-zero only when $(k-i) = \pm 1$. Using this information, (6.25) reduces to

$$y_k = a_k g_c(0) - b_{k-1} g_s(T_c) - b_{k+1} g_s(-T_c)$$

$$= a_k + \frac{1}{2}(-b_{k-1} + b_{k+1}). \qquad (6.26)$$

Using a similar derivation for z_k gives

$$z_k = b_k + \frac{1}{2}(a_{k-1} - a_{k+1}). \qquad (6.27)$$

Equations (6.26) and (6.27) have the form of partial response signals, with y_k and z_k having possible levels of 0, ±1 and ±2 as shown by the system eye-pattern for y(t) or z(t) in Fig. 6.4. Substituting (6.26) and (6.27) into (6.23) and (6.24), respectively, gives the two inputs to the decoder as

$$y_k' = a_k + \frac{1}{2}(-b_{k-1} + b_{k+1}) + n_{y_k} \qquad (6.28)$$

and

$$z_k' = b_k + \frac{1}{2}(a_{k-1} - a_{k+1}) + n_{z_k}. \qquad (6.29)$$

The following sections look at methods of "unwinding" this ISI.

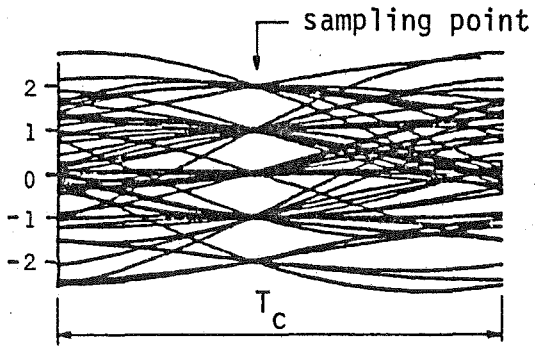


Fig. 6.4 Eye-pattern for either received channel of QVSB system when $\rho = 1$, showing 5 possible levels of the sampled signal.

6.4 PRECODING

The received signals in a QVSB system have been shown to have the form of PRS, where each sample at the input to the decoder is a weighted sum of several data bits input to the transmitter. The similarity between QVSB and PRS systems enables techniques developed to "unwind" the ISI in PRS systems to also be applied to QVSB. The simplest of these techniques is precoding which was developed by Lender [1963, 1964, 1966].

As the name implies, the precoding method incorporates the precoder stage in Fig. 6.1. Precoding is used to allow bit-by-bit decoding and avoids propagation of errors due to correlations among decoder input samples. The desired operation of the precoder is to relate the received levels to the input levels as shown in Table 6.1. This relationship can be described by the functions

$$\{(a_k + 1)/2\}_{\text{mod } 2} = (y_k)_{\text{mod } 2} \qquad (6.30)$$

and

$$\{(b_k + 1)/2\}_{\text{mod } 2} = (z_k)_{\text{mod } 2}. \tag{6.31}$$

TABLE 6.1

DESIRED INPUT-TO-RECEIVED LEVEL RELATIONSHIPS FOR QVSB WITH PRECODING

Input Levels		Received Levels	
a_k	b_k	y_k	z_k
+1	-	±1	-
-1	-	0 or ±2	-
-	+1	-	±1
-	-1	-	0 or ±2

To facilitate the use of logical processing, the precoding stage will be of the form shown in Fig. 6.5. Bilevel inputs a_k and b_k are converted to logical levels c_k and d_k prior to processing to logical levels c_k' and d_k'. A conversion back to bilevel signals a_k' and b_k' is required before modulation.

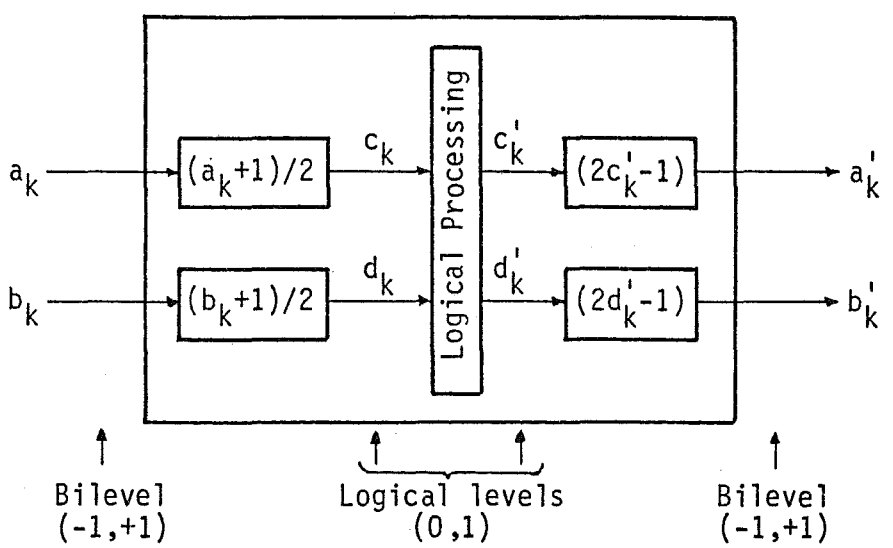


Fig. 6.5 Precoding stage.

The form of the logical processing can be derived as follows. From (6.30), (6.31) and Fig. 6.5,

$$(c_k)_{\text{mod } 2} = (y_k)_{\text{mod } 2} \tag{6.32}$$

and

$$(d_k)_{\text{mod } 2} = (z_k)_{\text{mod } 2}. \tag{6.33}$$

With the precoding stage, a'_k and b'_k replace, respectively, a_k and b_k in both (6.26) and (6.27). These equations can now be substituted into (6.32) and (6.33), giving

$$(c_k)_{\text{mod } 2} = \{a'_k + \tfrac{1}{2}(-b'_{k-1} + b'_{k+1})\}_{\text{mod } 2} \tag{6.34}$$

and

$$(d_k)_{\text{mod } 2} = \{b'_k + \tfrac{1}{2}(a'_{k-1} - a'_{k+1})\}_{\text{mod } 2}. \tag{6.35}$$

With the precoder's logical-to-bilevel conversion stage, (6.34) and (6.35) become, respectively,

$$(c_k)_{\text{mod } 2} = (2c'_k - 1 - d'_{k-1} + d'_{k+1})_{\text{mod } 2} \tag{6.36}$$

and

$$(d_k)_{\text{mod } 2} = (2d'_k - 1 + c'_{k-1} - c'_{k+1})_{\text{mod } 2}. \tag{6.37}$$

The terms $2c'_k$ in (6.36) and $2d'_k$ in (6.37) can be ignored since they will have values 0 or 2 which will not affect the modulo-2 summations. Modulo-2 addition and subtraction of the remaining logical levels can be performed by the exclusive-OR operator $\oplus$. Equations (6.36) and (6.37) can therefore be rewritten as, respectively,

$$c_k = 1 \oplus d'_{k-1} \oplus d'_{k+1} \tag{6.38}$$

and

$$d_k = 1 \oplus c'_{k-1} \oplus c'_{k+1}. \tag{6.39}$$

Rearranging (6.38) and (6.39) gives, respectively,

$$d'_{k+1} = c_k \oplus \overline{d}'_{k-1} \tag{6.40}$$

and

$$c'_{k+1} = d_k \oplus \overline{c}'_{k-1}. \tag{6.41}$$

These can be further simplified to

$$d'_k = c_k \oplus \overline{d}'_{k-2} \tag{6.42}$$

and

$$c'_k = d_k \oplus \overline{c}'_{k-2}, \tag{6.43}$$

which changes the relative input-output timing of the signals but not the modulo-2 decoding.

A computer simulation of the model shown in Fig. 6.1 was used to determine the QVSB system's BER performance with precoding and modulo-2 decoding. The results are plotted in Fig. 6.10, curve PREC.

6.5 PROBABILISTIC DECODING

While the precoding technique enables a simple decoder to be used, it makes no use of the inherent redundancy in the received signal. In this section the Viterbi algorithm [Viterbi, 1967], originally devised

for convolutional codes, which was shown by Kobayashi [Jan., Sept. and Dec. 1971] to have application in PRS systems, is used to improve the BER performance of the QVSB system.

6.5.1 Ignoring Noise Correlations

Matched filtering optimises SNR prior to decoding but, as shown in Appendix 4, results in correlations among the sampled noise levels. In this section a probabilistic decoding algorithm is derived which ignores the noise correlations; i.e., the noise samples are assumed to be statistically independent. The decoder is therefore non-optimum in terms of BER performance, but does give a marked improvement over the precoding technique.

The 2-dimensional probabilistic decoder [Wozencraft and Jacobs, 1965] is a simple extension of the 1-dimension decoder derived in section 2.3.4. The decoder sets estimates $(\hat{\underline{y}}, \hat{\underline{z}})$ equal to the hypothesis $(\underline{y}, \underline{z})$ which minimises the function

$$J(\underline{y}, \underline{z}) = \sum_{k=0}^{N} \left[(y'_k - y_k)^2 + (z'_k - z_k)^2 \right] \tag{6.44a}$$

$$= \sum_{k=0}^{N} \left[y'^2_k - 2y'_k y_k + y_k^2 + z'^2_k - 2z'_k z_k + z_k^2 \right]. \tag{6.44b}$$

The terms y'^2_k and z'^2_k in (6.44b) will remain fixed and independent of the alternative hypotheses $(\underline{y}, \underline{z})$ tested and can therefore be ignored when minimising $J(\underline{y}, \underline{z})$. Noting from (6.26) and (6.27) that signal levels $(y_0, z_1, y_2, z_3, \ldots.)$ are due to data bits $(a_0, b_1, a_2, b_3, \ldots.)$, and $(z_0, y_1, z_2, y_3, \ldots.)$ are due to $(b_0, a_1, b_3, a_4, \ldots.)$, then minimising $J(\underline{y}, \underline{z})$ is equivalent to independently maximising $J1(\underline{y}, \underline{z})$ and $J2(\underline{y}, \underline{z})$, where,

$$J1(\underline{y}, \underline{z}) = \sum_{\substack{k=0 \\ (k \text{ even})}}^{N} (2y_k' y_k - y_k^2) + \sum_{\substack{k=0 \\ (k \text{ odd})}}^{N} (2z_k' z_k - z_k^2) \quad (6.45)$$

and

$$J2(\underline{y}, \underline{z}) = \sum_{\substack{k=0 \\ (k \text{ odd})}}^{N} (2y_k' y_k - y_k^2) + \sum_{\substack{k=0 \\ (k \text{ even})}}^{N} (2z_k' z_k - z_k^2). \quad (6.46)$$

Considering only (6.45), since (6.46) has similar form, let $J1_\ell(\underline{y}, \underline{z})$ be the partial sum of the first ℓ terms in $J1(\underline{y}, \underline{z})$, then

$$J1_\ell(\underline{y}, \underline{z}) = \sum_{\substack{k=0 \\ (k \text{ even})}}^{\ell} (2y_k' y_k - y_k^2) + \sum_{\substack{k=0 \\ (k \text{ odd})}}^{\ell} (2z_k' z_k - z_k^2)$$

$$= \begin{cases} J1_{\ell-1}(\underline{y}, \underline{z}) + (2y_\ell' y_\ell - y_\ell^2), & \ell \text{ even} \\ J1_{\ell-1}(\underline{y}, \underline{z}) + (2z_\ell' z_\ell - z_\ell^2), & \ell \text{ odd.} \end{cases} \quad (6.47)$$

Equation (6.47) is in a recursive form used by the Viterbi decoder. The decoding can be performed by means of a trellis diagram [Viterbi, 1971] shown in Fig. 6.6. Each state is defined by the oldest two bits which contribute to a signal level y_ℓ or z_ℓ, and the transition to a new state by the latest bit; for example, state ($a_{\ell-1} = -1$, $b_\ell = 1$) can branch to ($b_\ell = 1$, $a_{\ell+1} = -1$) or ($b_\ell = 1$, $a_{\ell+1} = 1$) with associated signal outputs $z_\ell = 1$ or $z_\ell = 0$, respectively. For this example the branch transitions are given a length, or metric, $(2z_\ell' z_\ell - z_\ell^2)$. Maximising $J1(\underline{y}, \underline{z})$ requires finding the path through the trellis with the greatest metric.

Computer simulation results of BER performance are shown in Fig. 6.10, curve UNCORR. In order to obtain output sequences $\hat{\underline{a}}$ and $\hat{\underline{b}}$, a

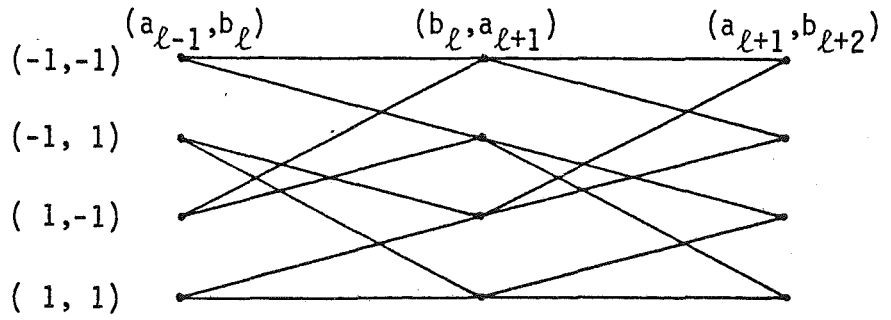


Fig. 6.6 Trellis representation of state transitions for probabilistic decoding of QVSB when noise correlations ignored.

truncated search length L, say, was used [Ungerboeck, 1974]. On calculation of the partial sum $J1_\ell(\underline{y}, \underline{z})$ to each of the four current states, the path history through only L previous states is retained. The oldest state in the path with the greatest metric is used for estimate $\hat{a}_{\ell-L}$ or $\hat{b}_{\ell-L}$. A value of L = 10 was found to be sufficient, where any further increase in the search length gave no significant improvement in BER.

6.5.2 Considering Noise Correlations

In this section a maximum-likelihood (ML) decoding algorithm which takes into account the noise statistics is derived for the QVSB system. This optimum decoder is seen to give an E_b/N_o improvement of approximately 0.8 dB over the non-optimum decoder derived in section 6.5.1.

The normalised autocorrelation and cross-correlation functions of the noise terms $n_y(t)$ and $n_z(t)$ are (see Appendix 4)

$$R'_{yy}(\tau) = R'_{zz}(\tau) = g_c(\tau) \tag{6.48}$$

and

$$R'_{zy}(\tau) = -R'_{yz}(\tau) = g_s(\tau). \tag{6.49}$$

As can be expected with matched filtering, these functions have identical form to the desired and interference data impulse response functions in Fig. 6.3.

ML decoding of $\underline{y}'$ and $\underline{z}'$ requires setting estimates $(\hat{\underline{y}}, \hat{\underline{z}})$ equal to the hypothesis $(\underline{y},\underline{z})$ which maximises the conditional probability density function [Wozencraft and Jacobs, 1965].

$$p(\underline{y}', \underline{z}' \mid \underline{y}, \underline{z}). \tag{6.50}$$

Using transformation of variables, maximising (6.50) is equivalent to maximising

$$p(\underline{y}' - \underline{y}, \underline{z}' - \underline{z} \mid y, z). \tag{6.51}$$

From (6.15) and (6.16), the terms $\underline{y}' - \underline{y}$ and $\underline{z}' - \underline{z}$ in (6.51) simply define the noise samples, which are independent of $\underline{y}$ and $\underline{z}$. Thus, (6.51) may be rewritten as

$$\begin{aligned} p(\underline{y}' - \underline{y}, \underline{z}' - \underline{z} \mid y, z) &= p(\underline{y}' - \underline{y}, \underline{z}' - \underline{z}) \\ &= p(n_{y_0}, n_{z_0}, n_{y_1}, n_{z_1}, \ldots, n_{y_N}, n_{z_N}) \\ &= p(\underline{n}), \text{ say.} \end{aligned} \tag{6.52}$$

From Wozencraft and Jacobs [1965]

$$p(\underline{n}) = \frac{\exp(-\tfrac{1}{2}\, \underline{n}\, \underline{C}^{-1}\, \underline{n}^{T})}{(2\pi)^{N+1}\, |\underline{C}|^{\frac{1}{2}}} \tag{6.53}$$

where,

$$\underline{C} = E[\underline{n}^{T}\, \underline{n}] \tag{6.54}$$

is the covariance matrix of $\underline{n}$.

Ungerboeck [1974] arrived at a likelihood function similar to (6.53), then used this to derive a "modified Viterbi algorithm" for general linear carrier-modulated data systems which encompasses the QVSB model. The "modified Viterbi algorithm" operates on the output of a matched filter where noise samples are correlated. In a more recent paper, Acampora [1981] derived a ML decoding algorithm for quadrature amplitude modulated signals where matched filtering and mutual interference between the quadrature related signals were considered. Both Ungerboeck and Acampora derived similar recursive relationships that enabled application of the Viterbi algorithm. An alternative approach, which makes specific use of the limited interference between the two VSB signals, will now be used in deriving a ML decoding algorithm for QVSB.

Noting, as in section 6.5.1, that the received signal ($\underline{y}'$, $\underline{z}'$) consists of two non-interacting data streams and that the correlations among noise samples, given by (6.48) and (6.49), follows a similar pattern, then (6.52) can be rewritten as

$$p(\underline{n}) = p(n_{y_0}, n_{z_1}, n_{y_2}, \ldots, n_{y_{(N-1)}}, n_{z_N}) \times p(n_{z_0}, n_{y_1}, n_{z_2}, \ldots$$

$$\ldots, n_{z_{(N-1)}}, n_{y_N}). \qquad (6.55)$$

Since independently maximising the two terms in (6.55) will require a similar operation, only the first term will be considered from this point on. This first term can be rewritten as

$$p(n_{y_0}, n_{z_1}, n_{y_2}, \ldots, n_{y_{(N-1)}}, n_{z_N}) = \prod_{\substack{k=0 \\ (k \text{ even})}}^{N-1} \left[p(n_{z_{(k+1)}} | n_{y_k}) p(n_{y_k} | n_{z_{(k-1)}}) \right]. \qquad (6.56)$$

From Wozencraft and Jacobs [1965],

$$p(n_{z_{(k+1)}}|n_{y_k}) = K_1 \exp\left[-(n_{z_{(k+1)}} - R'_{zy}(T_c)n_{y_k})^2/K_2\right] \qquad (6.57)$$

and

$$p(n_{y_k}|n_{z_{(k-1)}}) = K_1 \exp\left[-(n_{y_k} - R'_{yz}(T_c)n_{z_{(k-1)}})^2/K_2\right], \qquad (6.58)$$

where K_1 and K_2 are constants. From (6.49) and Fig. 6.3, $R'_{zy}(T_c) = -R'_{yz}(T_c) = \frac{1}{2}$. Substituting (6.57) and (6.58) into (6.56) and using the log-likelihood representation (see section 2.3.4), gives

$$L1(\underline{y}, \underline{z}) = \sum_{\substack{k=0 \\ (k \text{ even})}}^{N} \lambda_k, \qquad (6.59)$$

where, after expanding the n_y and n_z terms,

$$\lambda_k = -(z'_{k+1} - z_{k+1} - y'_k/2 + y_k/2)^2 - (y'_k - y_k + z'_{k-1}/2 - z_{k-1}/2)^2. \qquad (6.60)$$

Maximising (6.56) is equivalent to maximising $L1(\underline{y}, \underline{z})$.

Similar to section 6.5.1, the decoding can be performed using a trellis representation. As shown in Fig. 6.7, the trellis will have 8 states, each defined by the oldest three bits $(a_{\ell-2}, b_{\ell-1}, a_\ell)$ which contribute to calculation of the branch metric λ_ℓ. The branch transitions to the next states are determined by the latest bits $(b_{\ell+1}, a_{\ell+2})$. The combination $(a_{\ell-2}, b_{\ell-1}, a_\ell, b_{\ell+1}, a_{\ell+2})$ gives values $(z_{\ell-1}, y_\ell, z_{\ell+1})$ from (6.26) and (6.27) for calculating λ_ℓ. The Viterbi algorithm can be applied by noting that calculation of λ_ℓ is independent of metric calculations for transition branches into previous states. Thus, maximising

(6.59) can be carried out using the recursive relationship

$$L1_{\ell}(\underline{y}, \underline{z}) = \sum_{\substack{k=0 \\ (k,\ell \text{ even})}}^{\ell} \lambda_k$$

$$= L1_{\ell-2}(\underline{y}, \underline{z}) + \lambda_{\ell}. \qquad (6.61)$$

The BER performance for this ML decoding operation is shown in Fig. 6.10, curve CORR. A truncated search length of L = 6 was sufficient in this case.

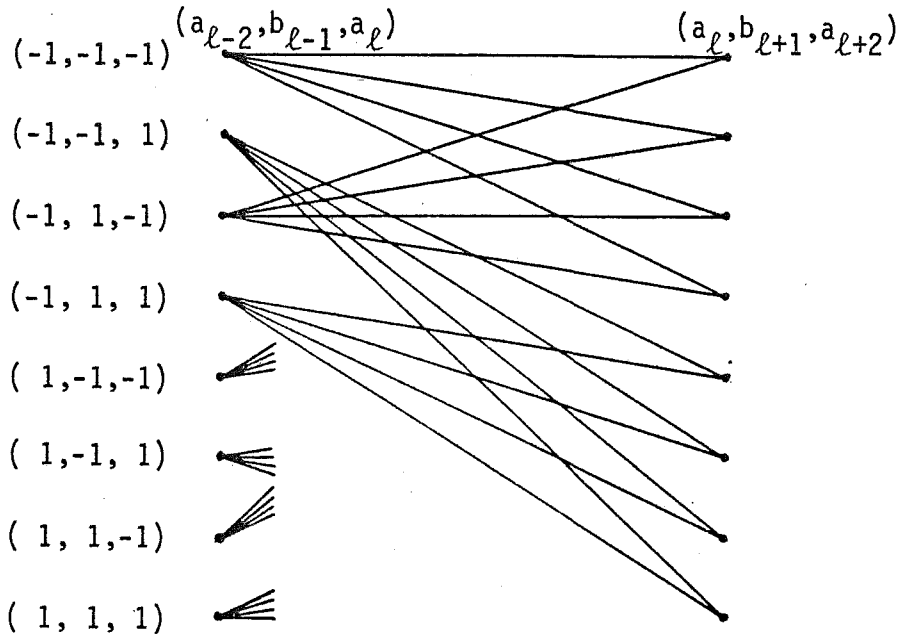


Fig. 6.7 Trellis representation of state transitions for probabilistic decoding of QVSB when noise correlations considered.

6.5.2.1 Simplifying metric computation

In computing branch metrics, the squaring operations in (6.60) could be accomplished by using a look-up table; however, such operations increase system complexity. A method to reduce this complexity, which does not appear to have received any previous consideration in the literature, is to replace squaring with an approximation using the absolute

value operation, whereby the metric calculation of (6.60) is simplified to

$$\lambda_k = -|z'_{k+1} - z_{k+1} - y'_k/2 + y_k| - |y'_k - y_k + z'_{k-1}/2 - z_{k-1}/2|. \tag{6.62}$$

The reason for considering such an approximation is that both squaring and absolute value functions belong to the general class of even ν^{th}-power-characteristics given by the relationship [Oberst and Schilling, 1971]

$$\xi = |\zeta|^{\nu}, \quad \nu > 0. \tag{6.63}$$

Assuming the use of binary notation, the divide-by-2 operations in (6.62) can easily be implemented by a shift-to-the-right. No multiplication or squaring operations are required when using the simplified metric. A performance degradation must be expected when using this approximation, but as shown in Fig. 6.10, curve CORRABS, only about 0.4 dB in E_b/N_o performance is conceded when (6.62) is used instead of (6.60).

6.5.3 Comparison of Processing Complexity

In addition to the relative BER performance of the probabilistic decoding algorithms, an important consideration in comparing their relative merits is the processing complexity required to implement the decoders. A comparison of this relative complexity, in terms of the computations required per unit time, is given in Table 6.2.

For each of the decoding algorithms two parallel and independent decoding operations are performed simultaneously. The computation indicated by (6.47), (6.60/6.61) and (6.61/6.62) must therefore be doubled. Also, the decoding algorithm where noise correlations are ignored requires

the latest set of branch metrics for each of the two independent decoding paths to be computed every T_c seconds. For both the optimum and simplified algorithms, where noise correlations are considered, branch metric computations for each decoding path are required only every $2T_c$ seconds.

TABLE 6.2

COMPARISON OF PROBABILISTIC DECODING ALGORITHMS

	DECODING ALGORITHMS		
	Ignoring Noise Correlations	Considering Noise Correlations	
		Optimum	Simplified
No. Summations/T_c sec	32	128	128
No. Binary Shifts/T_c sec	16	64	64
No. Multiplications or Squarings/T_c sec	32	32	0
No. Comparisons/T_c sec	8	8	8
Relative E_b/N_o Degradation at $P_b = 10^{-5}$ (dB)	0.8	0	0.4

6.5.4 Spectral Efficiency and Narrowing the QVSB Passband

Analysis thus far has assumed the shaping of the QVSB signal by the cascade of the transmitter and receiver bandpass filters to be that of a raised-cosine passband with rolloff factor $\rho = 1$. With optimum partitioning of the filtering between the transmitter and receiver the normalised spectral density of the QVSB signal in the channel is shown in Fig. 6.8. Also illustrated in Fig. 6.8 are spectral density plots for the cases where the overall raised-cosine passband has increased rolloff with $\rho = 0.75$ and $\rho = 0.5$.

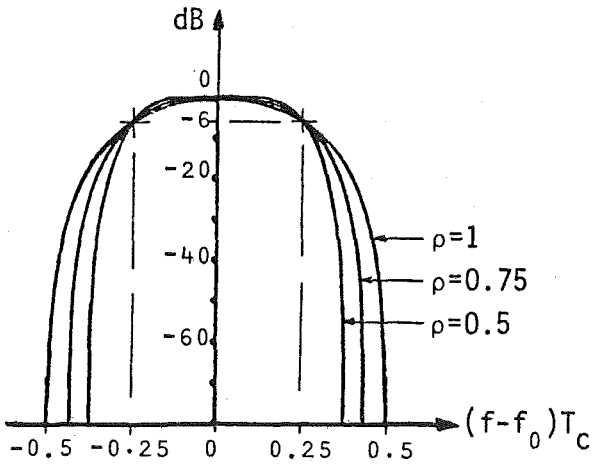


Fig. 6.8 Normalised spectral density plots for QVSB signal when overall signal shaping is a raised-cosine passband with indicated rolloff factor.

The half-power (-6 dB) bandwidth of the QVSB signal in the channel remains fixed at $1/(2T_c)$ Hz, independent of ρ. Using this as a definition of bandwidth results in a measure of spectral efficiency of 4 bits/sec/Hz, equivalent to 16-ary APK with Nyquist passband shaping. Note that $1/(2T)$ Hz is also the noise bandwidth of the QVSB receiver filter, independent of ρ.

Alternatively, using the first zero crossing bandwidth in defining spectral efficiency gives QVSB, with $\rho = 1$, a spectral efficiency of 2 bits/sec/Hz. This is equivalent to unfiltered 16-ary APK. Reducing ρ to 0.75 and 0.5 for the QVSB passband increases its spectral efficiency to 2.3 and 2.7 bits/sec/Hz, respectively.

For $\rho < 1$, the term $g_s((k-i)T_c)$ in (6.25) will have non-zero values for $(k-i) = \pm 1, \pm 3, \pm 5, \ldots$ in ever decreasing amplitudes. This is illustrated in Fig. 6.9 for the case where $\rho = 0.5$.

Since the complexity of the optimum decoder will grow exponentially with the number of received samples which are correlated, the non-optimum detection of QVSB for $\rho < 1$, in particular $\rho = 0.75$ and

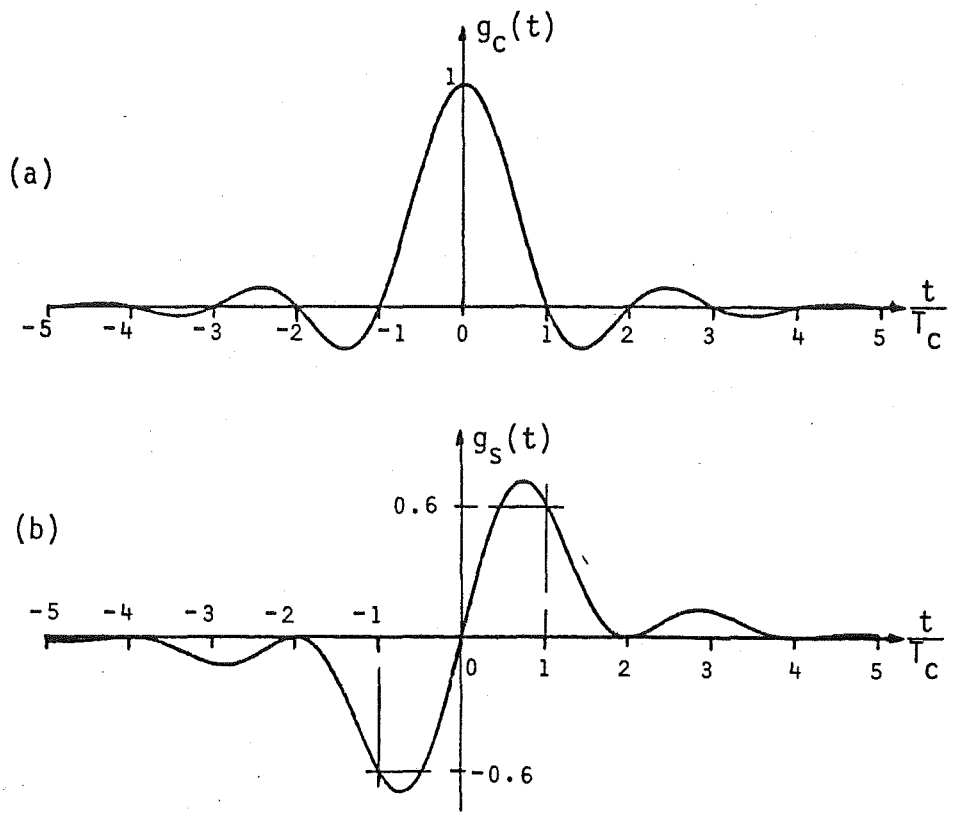


Fig. 6.9 (a) Desired and (b) interference data responses in QVSB system when $\rho = 0.5$.

$\rho = 0.5$, is considered. The decoding algorithm derived in section 6.5.2, equations (6.60) and (6.61), is used. The decoder is optimum only for the case where $\rho = 1$; it has knowledge of the interference due to $g_s(t)$ at times $t = \pm T_c$, so interference due to $g_s(t)$ being non-zero at $t = \pm 3T_c$, $\pm 5T_c$, $\pm 7T_c$, ... will appear as noise. For the two cases of $\rho = 0.75$ and $\rho = 0.5$, $g_s(T_c)$ is now, respectively, 0.56 and 0.6, so the y and z values in computing branch metrics must be changed accordingly; for example, when $\rho = 0.5$,

$$y_k = a_k + 0.6(-b_{k-1} + b_{k+1}) \tag{6.64}$$

and

$$z_k = b_k + 0.6(a_{k-1} - a_{k+1}). \tag{6.65}$$

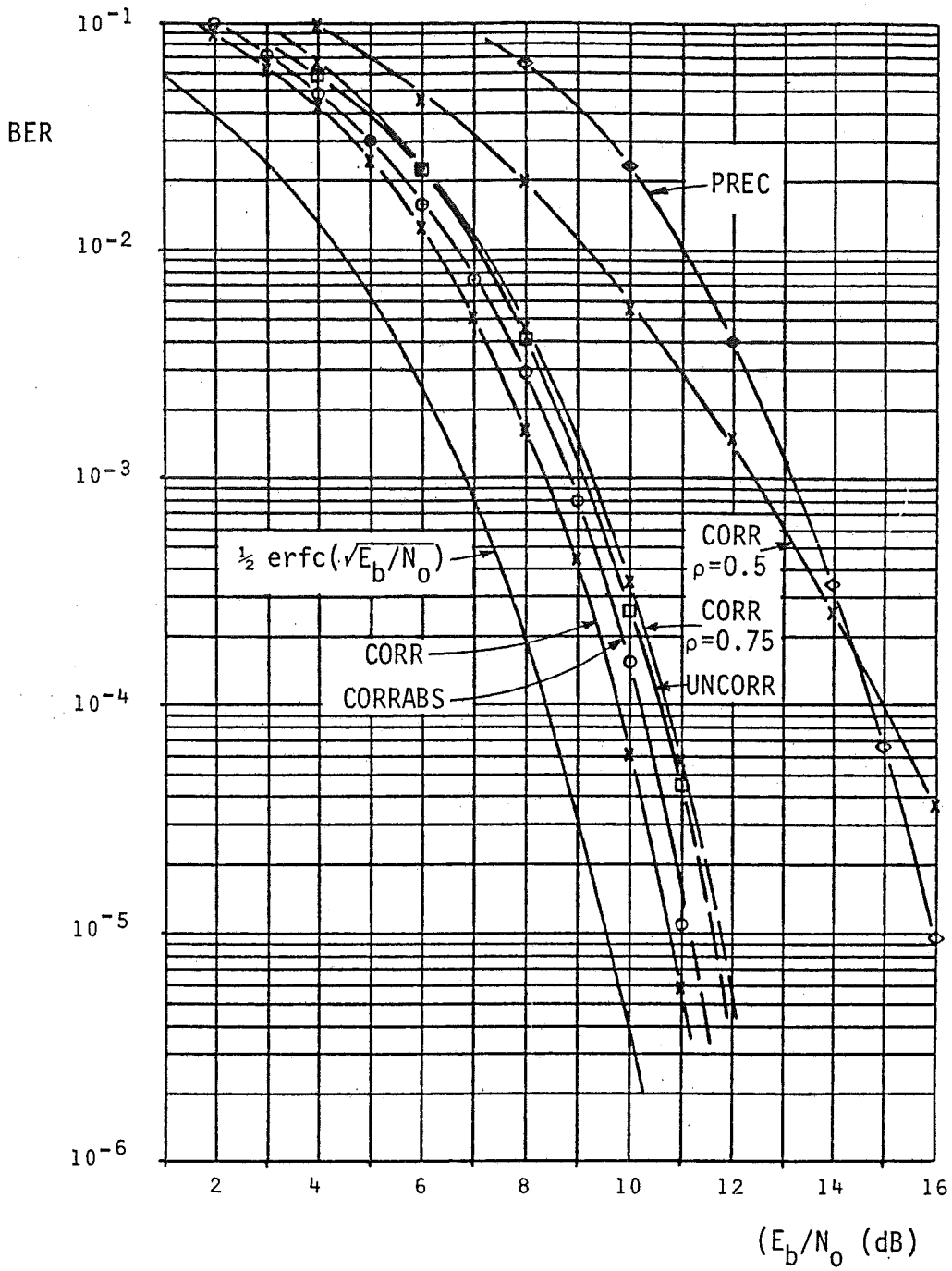


Fig. 6.10 Plots of simulation results. Unless otherwise stated, $\rho = 1$ for QVSB passband.

PREC - Precoding with modulo-2 decoding.
UNCORR - Viterbi algorithm when noise assumed uncorrelated.
CORR - Viterbi algorithm when noise correlations considered. For $\rho < 1$ decoding is non-optimum.
CORRABS - Viterbi algorithm when noise correlations considered, but using the absolute value approximation.

Simulation results for the two cases where $\rho = 0.75$ and $\rho = 0.5$ are shown in Fig. 6.10, curves CORR-$\rho = 0.75$ and CORR-$\rho = 0.5$, respectively. For $\rho = 0.75$ the E_b/N_o degradation for increasing the spectral efficiency without optimising the decoder is approximately 1 dB. For $\rho = 0.5$ the non-optimum decoder is unable to cope with the additional noise due to unaccounted for ISI. A marked degradation is apparent with this latter case as E_b/N_o increases and the ISI dominates the overall interference which is present.

6.6 QVSB AND QPSK AS IDENTITIES

A VSB data signal when considered with respect to its passband center frequency has the form of an OQPSK signal (see section 3.5). A similar identity can be derived for QVSB.

Ignoring the AWGN, the received QVSB signal is given by (see (6.12))

$$r(t) = \sqrt{2}\ \mathrm{Re}\left[\left\{(a'(t) - j\ b'(t)) * \left(g(t)e^{-j\frac{\pi t}{2T_c}}\right)\right\}e^{j\omega_1 t}\right]. \qquad (6.66)$$

The precoding stage in Fig. 6.1 is incorporated in deriving the identity, hence $a'(t)$ and $b'(t)$ are used in (6.66); their relationships with the precoder inputs $a(t)$ and $b(t)$ will be derived.

Equ. (6.66) defines $r(t)$ relative to the QVSB carrier frequency at the side of the passband. Alternatively, $r(t)$ can be defined relative to $f_0 = f_1 - 1/(4T)$, the passband center frequency, whereby

$$r(t) = \sqrt{2}\,\mathrm{Re}\left[\left\{(a'(t) - j\,b'(t)) * \left(g(t)e^{-j\frac{\pi t}{2T_c}}\right)\right\}e^{j\frac{\pi t}{2T_c}}\,e^{j\omega_0 t}\right]$$

$$= \sqrt{2}\,\mathrm{Re}\left[\left\{\left((a'(t) - j\,b'(t))e^{j\frac{\pi t}{2T_c}}\right) * g(t)\right\}e^{j\omega_0 t}\right]. \qquad (6.67)$$

Equ. (6.67) becomes clearer by including the precoding stage in the QVSB model and defining its input-output relationship as

$$a(t) - j\,b(t) = \{a'(t) - j\,b'(t)\}e^{j\frac{\pi t}{2T_c}}, \qquad (6.68a)$$

i.e.,

$$a'_k = \begin{cases} a_k(-1)^{k/2}, & k \text{ even} \\ b_k(-1)^{(k+1)/2}, & k \text{ odd} \end{cases} \qquad (6.68b)$$

and

$$b'_k = \begin{cases} b_k(-1)^{k/2}, & k \text{ even} \\ a_k(-1)^{(k-1)/2}, & k \text{ odd.} \end{cases} \qquad (6.68c)$$

Substituting (6.68a) into (6.67) gives

$$r(t) = \sqrt{2}\,\mathrm{Re}\left[\left\{(a(t) - j\,b(t)) * g(t)\right\}e^{j\omega_0 t}\right], \qquad (6.69)$$

which is the form of a QPSK signal with inputs a(t) and b(t) and passband shaping determined by g(t).

For the situation where h(t) is a raised-cosine bandpass filter with rolloff factor $\rho = 1$ (see Fig. 6.2), g(t) will be a lowpass filter with raised-cosine shaping as shown in Fig. 6.11. This is the passband

shaping function of a class-2 PRS system [Kretzmer, 1966], where the system polynomial is $1 + 2D + D^2$. Thus, with simple precoding and raised-cosine passband shaping with $\rho = 1$, the QVSB system is equivalent to QPSK modulated by $1 + 2D + D^2$ PRS signals on each quadrature channel.

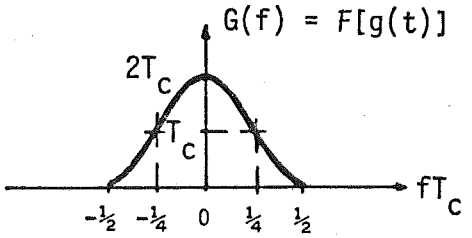


Fig. 6.11 Passband shaping function when QVSB considered relative to its passband center frequency.

6.7 GENERAL DISCUSSION AND COMPARISON OF RESULTS

It has been shown QVSB could be considered as a member of the PRS systems. QVSB, however, differs from the general form of PRS systems which have previously been developed. Rather than restrict the ISI to one phase plane, it is the mutual interference between the quadrature related VSB signals which introduces the ISI in the QVSB system.

The simple approach of precoding at the transmitter, with modulo-2 decoding at the receiver, exemplifies the similarity between QVSB and other PRS systems. The E_b/N_o degradation of 6.3 dB at $P_b = 10^{-5}$, relative to BPSK, also occurs for the PRS system with polynomial $1 + 2D - D^2$ [Huang and Feher, 1977; Kabal and Pasupathy, 1975]. This has ISI similar to QVSB, but contained within the one phase plane.

When QVSB is defined relative to its passband center frequency it is seen to have the form of a QPSK signal. If the QVSB passband has raised-cosine shaping with unity rolloff, then with the inclusion of a simple precoding stage, the system is identical to a QPSK signal modulated

by $1 + 2D + D^2$ PRS signals on each quadrature channel. Bit-by-bit detection of this class of PRS system also results in a 6.3 dB E_b/N_o degradation relative to BPSK [Kabal and Pasupathy, 1975].

The limited extent of ISI in the QVSB system makes probabilistic decoding an attractive solution to reducing its BER performance. Matched filtering results in correlations among noise samples input to the decoder. Ignoring these correlations enables the derivation of a non-optimum probabilistic decoder which gives a marked improvement over bit-by-bit detection. Using the Viterbi algorithm to implement the decoder gave simulation results which indicate an E_b/N_o performance improvement of 4.3 dB, at $P_b = 10^{-5}$, over the bit-by-bit detection method. A further improvement of 0.8 dB can be obtained by using a ML decoder that takes account of the noise correlation statistics. This improvement over the non-optimum probabilistic decoder is at the expense of a 4-fold increase in the number of summations and binary shifts required per T_c seconds in computing path metrics. The main computational complexity would, however, be due to the number of squaring and multiplication operations required per T_c seconds, and this is identical for both probabilistic decoding algorithms. This complexity can be removed from the optimum decoder, at a sacrifice of only 0.4 dB in E_b/N_o performance, by replacing squaring with absolute value operations. The use of this approximation may have application in other probabilistic decoding algorithms where, rather than expanding log-likelihood functions such as (6.44a), the squaring is replaced by the absolute value operation.

The spectral efficiency of QVSB is similar to that of 16-ary APK. Comparing ideal BER curves gives QVSB, with $\rho = 1$, an E_b/N_o advantage over 16-ary APK of 2.6 dB [Oetting, 1979]. As the QVSB passband is made more narrow the non-optimum detection of the received signal results in increased performance degradation. When $\rho = 0.75$ the extent of

unaccounted for ISI is relatively small and the non-optimum use of the decoding algorithm, derived for the case where $\rho = 1$, results in a 1 dB E_b/N_o performance degradation. When ρ is reduced further to 0.5 the non-optimum decoder's performance is degraded significantly due to the presence of unaccounted for ISI.

CHAPTER 7

CONCLUDING REMARKS AND RECOMMENDATIONS

7.1 GENERAL CONCLUSION

Several aspects of the digital transmission systems have been covered. The research has concentrated on spectrally efficient techniques, with particular emphasis on VSB as one such method. Alternatives to those systems presented could offer a greater spectral efficiency, however, an important consideration always retained is that spectral efficiency is only one of a number of important system parameters. The requirement for efficiency in power, complexity and cost must also be taken into account; no one parameter can be considered in isolation.

The baseband antipodal signalling model was defined and considered as a generally accepted datum for gauging the spectral and power efficiencies of other digital transmission systems. In the ideal, VSB and baseband antipodal signalling have identical P_b versus E_b/N_o curves and spectral efficiencies of 2 bits/sec/Hz. Of course, in practice, a degradation from the ideal must be accepted for both systems.

In Chapter 3 several alternatives to the conventional technique for generating and detecting VSB signals were given. The conventional approach is difficult if accurate bandpass filter shaping has to be defined, especially if high carrier frequency-to-data rate ratios are involved. The alternatives given enable passband shaping to be defined at or near baseband frequencies, thereby simplifying implementation.

The desire for simplicity in system implementation while retaining efficiencies in other areas was also maintained in the proposed VSB synchronization scheme in Chapter 4. Use of the synchronization pulses

gives a simple method of adding pilots to the VSB signal. Also, conventional circuitry is required at the receiver for recovering the carrier and bit timing waveforms.

In Chapter 5 it was shown that a signalling format with a constant envelope is not a necessary prerequisite for use over nonlinear channels when spectral efficiency is important. Previous research in this area has considered signals with a truncated time response. VSB (and OQPSK) with a truncated frequency response (cosine rolloff in this case) also appears to be a suitable signalling format. The VSB system compares favourably in terms of spectral efficiency and BER performance with alternative systems.

A point noted in Chapter 2 was the compromise that must be accepted among system parameters of spectral efficiency, power efficiency and complexity. This was the situation with the QVSB system analysed in Chapter 6. The spectral efficiency of VSB was doubled by using quadrature related carriers to VSB modulate two data signals into the one passband. Using the simple precoding technique to "unwind" the resulting ISI meant a marked E_b/N_o performance degradation had to be accepted compared with the conventional VSB system. Increasing the "intelligence" of the QVSB system by using non-optimum probabilistic decoding retrieved a large portion of the E_b/N_o performance lost in using the precoding technique, but at the expense of increased complexity. A further improvement, again requiring an increase in complexity, was obtained by a ML decoder. Replacing squaring with absolute value operations in this latter decoder simplified implementation of metric computation, but with an inevitable, though small, E_b/N_o performance degradation.

7.2 RECOMMENDATIONS FOR FUTURE WORK

At the conclusion of Chapter 4 on synchronizing VSB, mention

was made of the possibility of increasing the "intelligence" of the receiver to counteract the distorting effect of adding synchronization pulses to the transmitted data sequence. In general, if the data sequence is random the worst-case sequence, in terms of synchronization, will not occur very often so the information in the pilot tones is put to use only intermittently. However, if the information in the pilots must be used regularly to maintain synchronization, an increase in their power may be justified. This, in turn, may be justification for increasing the "intelligence" of the receiver so the distorting effect of the pilots can be negated by the decoder. If this is researched further, a way of training the decoder must be devised for it to synchronize itself with the times the pilots distort the recovered signal. One solution may be to initially transmit only synchronization pulses, possibly at full signal power, so enabling the decoder to recognise which samples at its input suffer a d.c. shift, and in which direction.

In the study of VSB signalling over a nonlinear channel, the transmitter and receiver filters were chosen on the basis that their cascade gave optimum filter shaping in a linear system; however, they may not necessarily hold as being optimum when cascaded with a nonlinear stage. Further research into passband shaping may result in a system with greater spectral and/or power efficiency.

The study of QVSB opens up several possibilities for future research. QVSB can be considered as one particular class of the PRS systems. A study of alternative forms of controlled ISI between quadrature related signals may lead to other, possibly more efficient, signalling formats. For the optimum decoder the problem of implementing squaring in the metric computations was simplified by using the absolute value operation instead. Use of this simplification may be worth considering for other probabilistic decoding algorithms, since results

indicate its use introduces only a small degradation in BER performance. A method of receiver synchronization must be studied before QVSB can be put into practice. Similar to QPSK, a fourth power device could be used to derive a local carrier at the receiver. For QVSB this method would require the transmitted data to be random so that, on average, signal power is always present in the DSB portion of the QVSB signal about its carrier frequency. In conjunction with this latter study, the effect of carrier phase errors and decoder sampling errors should also be considered.

REFERENCES

Acampora, A.S., "Analysis of maximum-likelihood sequence estimation performance for quadrature amplitude modulation," Bell Sys. Tech. J., vol. 60, pp. 865-885, July-Aug. 1981.

Anderson, C. and S. Barber, "Modulation : the key to effective bandwidth utilization," Telesis, pp. 172-179, Dec. 1979.

Andry, J.-P., G. de Passoz and J. Wajih, "80 kbit/s digital subscriber connection equipment over a 2-wire line," Commutation and Transmission, No. 3, pp. 61-68, 1981.

Austin, M.C. and M.U. Chang, "Quadrature overlapped raised-cosine modulation" IEEE Trans. Commun., vol. COM-29, pp. 237-249, Mar. 1981.

Becker, F.K., J.R. Davey and B.R. Saltzberg, "An AM vestigial sideband data transmission set using synchronous detection for serial transmission up to 3,000 bits per second," Trans. AIEE, pp. 97-101, May 1962.

Bellamy, J., Digital Telephony, John Wiley & Sons, 1982.

Bennet, W.R. and J.R. Davey, Data Transmission, McGraw-Hill, 1965.

Berlekamp, E.R., "The technology of error-correcting codes," Proc. IEEE, vol. 68, pp. 564-593, May 1980.

Bhargava, V.K., D. Haccoun, R. Matyas and P.P. Nuspl, Digital Communications by Satellite, John Wiley and Sons, 1981.

Bowman, D.F., P. Bylanski and J. Tritton, "Studies of subscriber's digital telephone terminal," Int. Zurich Seminar on Digital Commun., pp. D3/1-5, Mar. 1978.

Buckrell, P.F., "Digitization of the telephone subscriber network," NELCON 80, pp. 93-100, 1980.

Bylanski, P. and D.G.W. Ingram, Digital Transmission Systems, Peter Peregrinus Ltd., 1976.

Bylanski, P., G.N. Lawrence and J.A. Tritton, "The digital telephone," GEC J. Sci. Tech., vol. 46, pp. 19-24, 1980.

Choquet, M.F. and H.J. Nussbaumer, "Generation of synchronous data transmission signals by digital echo modulation," IBM J. Res. Develop., pp. 364-377, Sept. 1971.

Clark, A.P., Principles of Digital Data Transmission, Pentech Press, 1976.

Croisier, A. and J.D. Pierret, "The digital echo modulation," IEEE Trans. Commun. Tech., vol. COM-18, pp. 367-376, Aug. 1970.

Davenport, W.B. Jr., and W.L. Root, Random Signals and Noise, McGraw-Hill, 1958.

Del Re, E., "A property of digital quadrature filters," Proc. IEEE, vol. 69, pp. 1577-1578, Dec. 1981.

Dorros, I., "Telephone nets go digital," IEEE Spectrum, pp. 48-53, Apr. 1983.

Fang, R.J.F. (May 1981), "Quaternary transmission over satellite channels with cascaded nonlinear elements and adjacent channel interference," IEEE Trans. Commun., vol. COM-29, pp. 567-581.

Fang, Y. (Apr. 1972), New methods of generating vestigial-sideband signal for data transmission," IEEE Trans. Commun., vol. COM-20, pp. 147-157.

Feher, K. (1981), Digital Communications : Microwave Applications, Prentice-Hall Inc.

Feher, K. (1983), "Digital modulation," Satellite Communications, pp. 36-41, Feb.

Forney, G.D. Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," IEEE

Trans. Info. Theory, vol. IT-18, pp. 362-378, May 1972.

Franks, L.E., "Carrier and bit synchronization in data communication - a tutorial review," IEEE Trans. Commun., vol. COM-28, pp. 1107-1121, Aug. 1980.

Franks, L.E. and J.P. Bubrouski, "Statistical properties of timing jitter in a PAM timing recovery scheme," IEEE Trans. Commun., vol. COM-22, pp. 913-920, July 1974.

Gardner, F.M., Phase Lock Techniques, John Wiley and Sons, 1979.

Gibby, R.A. and J.W. Smith, "Some extensions of Nyquist's telegraph transmission theory," Bell Sys. Tech. J., pp. 1487-1510, Sept. 1965.

Gronemeyer, S.A. and A.L. McBride, "MSK and offset QPSK modulation," IEEE Trans. Commun., vol. COM-24, pp. 809-819, Aug. 1976.

Haykin, S., Communication Systems, John Wiley and Sons, 1978.

Heller, J.A., "Viterbi decoding for satellite and space communication," IEEE Trans. Commun., vol. COM-19, pp. 835-848, Oct. 1971.

Henderson, B.G. and J.A. Webb (Jan. 1982), "A bandwidth-efficient data transmission system," Proc. Pacific Telecom. Conf., pp. 121-122.

Henderson, B.G. and J.A. Webb (1984), "Synchronizing VSB data transmission," J. Electrical and Electronic Engineers, Australia. To be published.

Hill, F.S. Jr., "Optimum pulse shapes for pulse-amplitude modulation data transmission using vestigial sideband modulation," IEEE Trans. Commun., vol. COM-23, pp. 352-361, Mar. 1975.

Ho, E.Y., "A new carrier recovery technique for vestigial sideband (VSB) data systems," IEEE Trans. Commun., vol. COM-22, pp. 1866-1870, Nov. 1974.

Holte, N. and S. Stuefflotten, "A new digital echo canceler for two-wire subscribers lines," IEEE Trans. Commun., vol. COM-29, pp. 1573-1581, Nov. 1981.

Huang, J. and K. Feher (1977), "On partial response transmission systems," Int. Conf. Commun., vol. 1., pp. 3.3/47-51.

Huang, J.C.Y., K. Feher and M. Gendron (Nov. 1979), "Techniques to generate ISI and jitter-free bandlimited Nyquist signals and a method to analyze jitter effects," IEEE Trans. Commun., vol. COM-27, pp. 1700-1711.

Inoue, N., R. Komiya and Y. Inoue, "Time-shared two-wire digital transmission for subscriber loops," IEEE Int. Conf. Commun., pp. 2.4/1-5, 1979.

Kabal, P. and S. Pasupathy, "Partial-response signaling," IEEE Trans. Commun., vol. COM-23, pp. 921-934, Sept. 1975.

Kalet, I. and S.B. Weinstein, "In-band generation of synchronous linear data signals," IEEE Trans. Commun., vol. COM-21, pp. 1116-1122, Oct. 1973.

Kelly, M.W., Echo Cancellation on Communication Circuits, Ph.D. Thesis, University of Canterbury, New Zealand, 1979.

Kennedy, D.J. and O. Shimbo, "Cochannel interference in nonlinear QPSK satellite systems," IEEE Trans. Commun., vol. COM-29, pp. 582-592, May 1981.

Kobayashi, H. (Jan. 1971), "Application of probabilistic decoding to digital magnetic recording systems," IBM J. Res. Develop., pp. 64-74.

Kobayashi, H. (Sept. 1971), "Correlative level coding and maximum-likelihood decoding," IEEE Trans. Info. Theory, vol. IT-17, pp. 586-594.

Kobayashi, H. (Dec. 1971), "A survey of coding schemes for transmission or recording of digital data," IEEE Trans. Commun. Tech., vol. COM-19, pp. 1087-1100.

Kobayashi, H. and D.T. Tang, "On decoding of correlative level coding systems with ambiguity zone detection," IEEE Trans. Commun. Tech., vol. COM-19, pp. 467-477, Aug. 1971.

Kretzmer, E.R., "Generalization of a technique for binary communication," IEEE Trans. Commun. Tech., pp. 67-68, Feb. 1966.

Lender, A. (May 1963), "The duobinary technique for high speed data transmission," IEEE Trans. Commun. Electron., vol. 82, pp. 214-218.

Lender, A. (Dec. 1964), "Correlative digital communication techniques," IEEE Trans. Commun. Tech., pp. 128-135.

Lender, A. (Feb. 1966), "Correlative level coding for binary-data transmission," IEEE Spectrum, vol. 3, pp. 104-115.

Le-Ngoc, L. and K. Feher (Feb. 1983), "Performance of an IJF-OQPSK modem in cascaded nonlinear and regenerative satellite systems," IEEE Trans. Commun., vol. COM-31, pp. 296-301.

Le-Ngoc, L., K. Feher and H. Pham Van (Jan. 1982), "New modulation techniques for low-cost power and bandwidth efficient satellite earth stations," IEEE Trans. Commun., vol. COM-30, pp. 275-283.

Lowe, S.J. and J.A. Webb, "Narrow band data transmission," Proc. IEEE Conf. on Circuits and Computers, 1980.

Lucky, R.W., J. Salz and E.J. Weldon, Jr., Principles of Data Communication, 1968.

Mathwich, R.H., J.F. Balcewicz and M. Hecht, "The effect of tandem band and amplitude limiting on the E_b/N_o performance of minimum (frequency) shift keying (MSK)," IEEE Trans. Commun., vol. COM-22, pp. 1525-1539, Oct. 1974.

Meyer, J. and T. Roste, "Field trials of two-wire digital transmission in the subscriber loop plant," Int. Conf. Commun., pp. 2.5/1-5, 1979.

Newcombe, E.A. and S. Pasupathy, "Error rate monitoring for digital communications," Proc. IEEE, vol. 70, pp. 805-828, Aug. 1982.

Nyquist, H., "Certain topics in telegraph transmission theory," Trans. AIEE, pp. 617-644, Feb. 1928.

NZPO (Oct. 1977), Subscribers line survey 1975-1976, Engineering Report/ Res. & Develop. 2067.

Oberst, J.F. and D.L. Schilling, "The SNR of a frequency doubler," IEEE Trans. Commun. Tech., pp. 97-99, Feb. 1971.

Oetting, J.D., "A comparison of modulation techniques for digital radio," IEEE Trans. Commun., vol. COM-27, pp. 1752-1762, Dec. 1979.

Panter, P.F., Modulation, Noise and Spectral Analysis, McGraw-Hill, 1965.

Parsons, J.D. and A.U.H. Sheikh, "The characteristics of impulsive noise and considerations for a noise-measuring receiver," The Radio and Electronics Engineer, vol. 49, pp. 467-476, Sept. 1979.

Pasupathy, S. (July 1979), "Minimum shift keying : a spectrally efficient modulation," IEEE Commun. Mag., pp. 14-22.

Pless, V., Introduction to the Theory of Error-Correcting Codes, John Wiley and Sons, 1982.

Rhodes, S.A., "Effects of hardlimiting on bandlimited transmission with conventional and offset QPSK modulation," Nat. Telecom. Conf., pp. 20F/1-7, 1972.

Rich, M.A., "Designing phase locked loop oscillators for synchronization," IEEE Trans. Commun., vol. COM-22, pp. 890-896, July 1974.

Ristenbatt, M.P., "Alternatives in digital communications," Proc. IEEE, vol. 61, pp. 703-721, June 1973.

Schwartz, J.W., J.M. Aein and J. Kaiser, "Modulation techniques for multiple access to a hard-limiting satellite repeater," Proc. IEEE, vol. 54, pp. 763-777, May 1966.

Shannon, C.E. (Oct. 1948), "A mathematical theory of communication," Bell Sys. Tech. J., vol. 27, pp. 379-423, July, and pp. 623-656.

Shannon, C.E. (Jan. 1949), "Communication in the presence of noise," Proc. IRE, vol. 37, pp. 10-21.

Smith, B.M., "Some results for the eye-openings of digital signals with raised-cosine frequency characteristics," A.T.R., vol. 16, pp. 16-22, 1982.

Strauss, R., J. Betting and R. Metivier, "Travelling wave tubes for communication satellites," Proc. IEEE, vol. 65, pp. 387-400, Mar. 1977.

Sunde, E.D. (May 1954), "Theoretical fundamentals of pulse transmission - I," Bell Sys. Tech. J., pp. 721-788.

Sunde, E.D. (July 1954), "Theoretical fundamentals of pulse transmission - II," Bell Sys. Tech. J., pp. 987-1010.

Svensson, T., "Methods of two-wire duplex transmission at 80 kbit/s on subscriber lines," Int. Zurich Seminar on Digital Commun., pp. C5/1-4, 1978.

Thomas, C.M., M.Y. Weidner and S.H. Durrani, "Digital amplitude-phase keying with M-ary alphabets," IEEE Trans. Commun., vol. COM-22, pp. 168-179, Feb. 1974.

Torbergsen, R. and J. Meyer, "A digital subscriber set," Int. Zurich Seminar on Digital Commun., pp. D2/1-4, Mar. 1978.

Ungerboeck, G., "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," IEEE Trans. Commun., vol. COM-22, pp. 624-636, May 1974.

van den Elzen, H.C., "On the theory and calculation of worst-case eye openings in data-transmission systems," Philips Research Reports, vol. 30, pp. 385-435, Dec. 1975.

Van Trees, H.L., E.V. Haversten and T.P. McGarty, "Communication satellites : looking to the 1980s," IEEE Spectrum, vol. 14, pp. 42-51, Dec. 1977.

Viterbi, A.J. (Apr. 1967), "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," IEEE Trans. Info. Theory, vol. IT-13, pp. 260-269.

Viterbi, A.J. (Oct. 1971), "Convolutional codes and their performance in communication systems," IEEE Trans. Commun. Tech., vol. COM-19, pp. 751-772.

Weaver, D.K. Jr., "A third method of generation and detection of single-sideband signals," Proc. IRE, pp. 1703-1705, Dec. 1956.

Weinstein, S.B., "Generation of linear and nonlinear synchronous data signals by digital echo modulation in a time-varying transversal filter," IEEE Trans. Commun., pp. 112-118, Jan. 1976.

Williams, M.B., "The characteristics of telephone circuits in relation to data transmission," Post Office Electrical Engineers' J., vol. 59, pp. 151-162, Oct. 1966.

Wozencraft, J.M. and I.M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, 1965.

APPENDIX 1

THEORETICAL LIMIT IN E_b/N_o FOR ERROR-FREE TRANSMISSION

(refer section 2.2.2)

The relationship between E_b/N_o and C/W, the theoretical spectral efficiency of a channel, is (see (2.8),)

$$E_b/N_o = \frac{2^{C/W} - 1}{C/W} \tag{A1.1}$$

For a fixed information rate C, the limiting value in E_b/N_o for error-free transmission is found when infinite bandwidth is available, i.e. when C/W = 0.

Let

$$f(C/W) = 2^{C/W}, \tag{A1.2}$$

then using the MacLaurin series expansion

$$f(x) = f(0) + x\frac{df(0)}{dx} + \frac{x^2}{2!}\frac{d^2f(0)}{dx^2} + \ldots, \tag{A1.3}$$

gives

$$f(C/W) = 1 + \frac{C}{W}\log_e 2 + \left(\frac{C}{W}\right)^2 \frac{\log_e^2 2}{2!} + \ldots \tag{A1.4}$$

Substituting (A1.4) into (A1.1) gives

$$E_b/N_o = \log_e 2 + \frac{\left(\frac{C}{W}\right)\log_e^2 2}{2!} + \frac{\left(\frac{C}{W}\right)^2 \log_e^3 2}{3!} + \ldots, \tag{A1.5}$$

so

$$\lim_{C/W \to 0} E_b/N_o = \log_e 2$$

$$= 0.693, \qquad \text{(A1.6)}$$

or in decibels

$$\lim_{C/W \to 0} E_b/N_o = 10 \log_{10}(0.693) \text{ dB}$$

$$= -1.6 \text{ dB}. \qquad \text{(A1.7)}$$

Thus, the theoretical limit in the value of E_b/N_o for error-free transmission is -1.6 dB.

APPENDIX 2

AUTOCORRELATION FUNCTION FOR A DATA SEQUENCE OF RECTANGULAR PULSES

(refer section 2.3.2)

Initially, consider the data sequence c(t) to consist of return-to-zero pulses as shown in Fig. A2.1, where

$$c(t) = \sum_{\ell} \frac{c_\ell}{T'} \text{rect}\left(\frac{t - \ell T_c}{T'}\right). \qquad \text{(A2.1)}$$

The pulses are assumed to be of random polarity and amplitude.

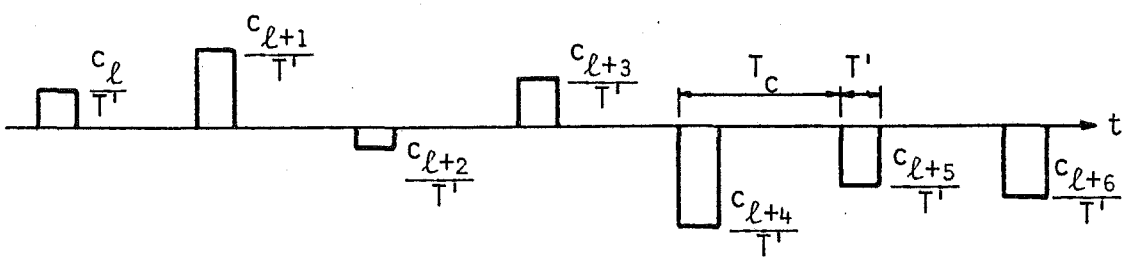


Fig. A2.1 Return-to-zero data sequence c(t).

As $T' \to 0$, an impulse sequence

$$\lim_{T' \to 0} c(t) = \sum_{\ell} c_\ell \, \delta(t - \ell T_c) \qquad \text{(A2.2)}$$

will be obtained.

The autocorrelation function of c(t) is

$$R_{cc}(\tau) = E[c(t + \tau)\, c(t)]. \qquad \text{(A2.3)}$$

For the case where $|\tau| > T'$ the data sequences $c(t + \tau)$ and $c(t)$ will be independent of each other, so

$$R_{cc}(\tau) = E[c(t + \tau)]\, E[c(t)]$$

$$= 0 \qquad , \ |\tau| > T' \tag{A2.4}$$

When $\tau = 0$,

$$R_{cc}(0) = E[c^2(t)]$$

$$= \frac{T'}{T_c} E\left[\left(\frac{c_\ell}{T'}\right)^2\right]. \tag{A2.5}$$

As $|\tau|$ increases from 0 to T', $R_{cc}(\tau)$ will be a linearly decreasing function, so in general

$$R_{cc}(\tau) = \begin{cases} \left(1 - \frac{|\tau|}{T'}\right) \frac{E[c_\ell^2]}{T_c\, T'} & , \quad |\tau| < T' \\ 0 & , \quad \text{elsewhere.} \end{cases} \tag{A2.6}$$

$R_{cc}(\tau)$ is shown in Fig. A2.2. As $T' \to 0$, $R_{cc}(\tau)$ will have the form of a delta function, where

$$\lim_{T' \to 0} R_{cc}(\tau) = \frac{E[c_\ell^2]}{T_c}\, \delta(\tau). \tag{A2.7}$$

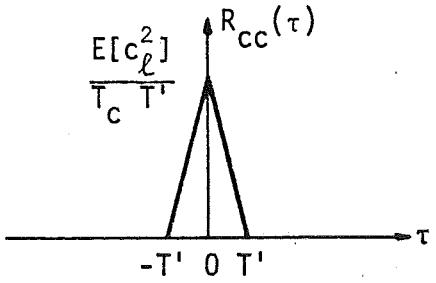


Fig. A2.2 Autocorrelation function of a random data sequence of rectangular pulses.

APPENDIX 3

CALCULATION OF σ_α^2 AND E_b FOR VSB

(refer section 3.2.2)

A3.1 CALCULATION OF σ_α^2

The spectral density of the noise $\beta(t)$ at the output of the receiver filter is

$$S_\beta(f) = \frac{N_o}{2} |H_2(f)|^2$$

$$= \frac{N_o}{2} H(f), \qquad \text{(A3.1)}$$

so the power of the noise at this point is

$$\sigma_\beta^2 = E[\beta^2(t)]$$

$$= \int_{-\infty}^{\infty} S_q(f)\, df$$

$$= \frac{N_o}{2} h(0) \qquad \text{(A3.2)}$$

Noting that half of the noise power at the output of the receiver mixer falls outside the passband of $\ell(t)$, the power of $\alpha(t)$ is therefore given by

$$\sigma_\alpha^2 = \tfrac{1}{2} E[\{\beta(t)\ \sqrt{2}\ \cos(\omega_1 t)\}^2]$$

$$= E[\beta^2(t)]\ E[\cos^2(\omega_1 t)]$$

$$= \tfrac{1}{2} E[\beta^2(t)]$$

$$= \frac{N_o\, h(0)}{4} . \qquad \text{(A3.3)}$$

Since

$$h(t) = R_e\left[\{h_c(t) + jh_s(t)\}e^{j\omega_1 t}\right], \qquad \text{(A3.4)}$$

then

$$h(0) = h_c(0), \qquad \text{(A3.5)}$$

so

$$\sigma_\alpha^2 = \frac{N_o\, h_c(0)}{4} . \qquad \text{(A3.6)}$$

A3.2 CALCULATION OF E_b

With reference to Fig. 3.1, let

$S_d(f)$ = spectral density of $d(t)$

$$= \frac{E[d_k^2]}{T}$$

$$= \frac{1}{T} , \qquad \text{(A3.7)}$$

$S_\ell(f)$ = spectral density of signal at output of transmitter lowpass filter

$$= S_d(f)\ |L(f)|^2 \qquad \text{(A3.8)}$$

and

$S_m(f)$ = spectral density of signal at output of transmitter mixer

$$= \tfrac{1}{2} S_\ell(f - f_1) + \tfrac{1}{2} S_\ell(f + f_1), \tag{A3.9}$$

then the spectral density of q(t) is

$$S_q(f) = S_m(f) \, |H_1(f)|^2$$

$$= \frac{1}{2T} \{|L(f - f_1)|^2 + |L(f + f_1)|^2\} \, |H_1(f)|^2. \tag{A3.10}$$

Since it has been assumed L(f) takes no part in signal shaping, then

$$S_q(f) = \frac{1}{2T} \, |H_1(f)|^2$$

$$= \frac{1}{2T} H(f), \tag{A3.11}$$

so

$$E_b = T \int_{-\infty}^{\infty} S_q(f) \, df$$

$$= \frac{h(0)}{2} . \tag{A3.12}$$

Using (A3.5) gives

$$E_b = \frac{h_c(0)}{2} . \tag{A3.13}$$

APPENDIX 4

AUTOCORRELATION AND CROSS-CORRELATION FUNCTIONS FOR BASEBAND NOISE IN QVSB SYSTEM

(refer section 6.5.2)

The noise at the output of the receiver bandpass filter is

$$\beta(t) = n(t) * h_2(t)$$

$$= n_c(t)\cos(\omega_o t) - n_s(t)\sin(\omega_o t), \qquad \text{(A4.1)}$$

where $n_c(t)$ and $n_s(t)$ are independent baseband Gaussian variables with the same power as $\beta(t)$. The spectral density of $\beta(t)$ is

$$S_\beta(f) = \frac{N_o}{2}|H_2(f)|^2$$

$$= \frac{N_o}{2}H(f), \qquad \text{(A4.2)}$$

where, (see (6.9))

$$h(t) = F[H(f)]$$

$$= 2g(t)\cos(\omega_o t). \qquad \text{(A4.3)}$$

The autocorrelation function of $\beta(t)$ is therefore

$$R_{\beta\beta}(\tau) = F[S_\beta(f)]$$

$$= N_o\, g(\tau)\cos(\omega_o \tau). \qquad \text{(A4.4)}$$

Haykin [1978] showed the spectral densities of $n_c(t)$ and $n_s(t)$ are given by

$$S_c(f) = S_s(f) = \begin{cases} S_\beta(f - f_o) + S_\beta(f + f_o), & |f| \leq B/2 \\ 0 & , \quad \text{elsewhere}, \end{cases} \tag{A4.5}$$

where B Hz is the bandwidth of the noise at the output of $h_2(t)$. Using (A4.2), (A4.3) and (A4.5) the autocorrelation functions of $n_c(t)$ and $n_s(t)$ are

$$R_{cc}(\tau) = R_{ss}(\tau) = \frac{N_o}{2} \int_{-B/2}^{B/2} \{H(f - f_o) + H(f + f_o)\} e^{j2\pi f\tau}\, d\tau$$

$$= N_o\, g(\tau). \tag{A4.6}$$

The baseband noise terms $n_y(t)$ and $n_z(t)$ are given in (6.19) and (6.20). The autocorrelation function of $n_y(t)$ is

$$R_{yy}(\tau) = E[n_y(t + \tau)\, n_y(t)]$$

$$= \tfrac{1}{2} E\left[\left\{n_c(t + \tau) \cos\left(\frac{\pi(t + \tau)}{2T_c}\right) + n_s(t + \tau) \sin\left(\frac{\pi(t + \tau)}{2T_c}\right)\right\}\right.$$

$$\left. \times \left\{n_c(t) \cos\left(\frac{\pi t}{2T_c}\right) + n_s(t) \sin\left(\frac{\pi t}{2T_c}\right)\right\}\right]. \tag{A4.7}$$

Noting that, since $n_c(t)$ and $n_s(t)$ are independent,

$$E[n_c(t_1)\, n_s(t_2)] = 0, \tag{A4.8}$$

and also

$$E\left[\cos\left(\frac{2\pi t + \pi\tau}{2T_c}\right)\right] = E\left[\sin\left(\frac{2\pi t + \pi\tau}{2T_c}\right)\right]$$

$$= 0, \qquad \text{(A4.9)}$$

then (A4.7) reduces to

$$R_{yy}(\tau) = \tfrac{1}{4} E[n_c(t + \tau)\, n_c(t)] \cos\left(\frac{\pi\tau}{2T_c}\right)$$

$$+ \tfrac{1}{4} E[n_s(t + \tau)\, n_s(t)] \cos\left(\frac{\pi\tau}{2T_c}\right)$$

$$= \tfrac{1}{4} [R_{cc}(\tau) + R_{ss}(\tau)] \cos\left(\frac{\pi\tau}{2T_c}\right)$$

$$= \frac{N_o}{2} g(\tau) \cos\left(\frac{\pi\tau}{2T_c}\right)$$

$$= \frac{N_o}{2} g_c(\tau). \qquad \text{(A4.10)}$$

Similarly, the autocorrelation function of $n_z(t)$ is given by

$$R_{zz}(\tau) = \frac{N_o}{2} g_c(\tau). \qquad \text{(A4.11)}$$

The cross-correlation function of $n_z(t)$ and $n_y(t)$ is

$$R_{zy}(t) = E[n_z(t + \tau)\, n_y(t)]$$

$$= \tfrac{1}{2} E\left[\left\{n_c(t + \tau) \sin\left(\frac{\pi(t + \tau)}{2T_c}\right) - n_s(t + \tau) \cos\left(\frac{\pi(t + \tau)}{2T_c}\right)\right\}\right.$$

$$\left. \times \left\{n_c(t) \cos\left(\frac{\pi t}{2T_c}\right) + n_s(t) \sin\left(\frac{\pi t}{2T_c}\right)\right\}\right]$$

$$= \tfrac{1}{4}\, E[n_c(t+\tau)n_c(t)]\, \sin\left(\frac{\pi t}{2T_c}\right)$$

$$+ \tfrac{1}{4}\, E[n_s(t+\tau)n_s(t)]\, \sin\left(\frac{\pi \tau}{2T_c}\right)$$

$$= \frac{N_o}{2}\, g(\tau)\, \sin\left(\frac{\pi \tau}{2T_c}\right)$$

$$= \frac{N_o}{2}\, g_s(\tau) \qquad \text{(A4.12)}$$

$$= -R_{yz}(\tau).$$

The normalised autocorrelation and cross-correlation functions of the baseband noise are therefore

$$R'_{yy}(\tau) = R'_{zz}(\tau) = R_{yy}(\tau)/R_{yy}(0)$$

$$= g_c(\tau), \qquad \text{(A4.13)}$$

and

$$R'_{zy}(\tau) = -R_{yz}(\tau) = R_{zy}(\tau)/\sqrt{R_{yy}(0)\, R_{zz}(0)}$$

$$= g_s(\tau). \qquad \text{(A4.14)}$$